

Twilight of forgotten dreams?

The plan to close Britain's first commercial nuclear plant should not be misread (as it has been) as a further nail in the coffin of the nuclear industry worldwide. But there is much to be done.

WHATEVER will happen to the nuclear power industry in the industrialized West? And, for that matter, in the Soviet Union? In the past few months, the United States has followed the precedent set by Austria in 1986 in deciding to dispense with the electricity that comes from two nuclear power plants, in New Hampshire and on Long Island. Now there is subdued rejoicing in Britain, among those who believe nuclear power to be an abomination, at the decision that the first commercial nuclear plant to be built, a gas-cooled reactor at Berkeley (pronounced "Barklee", contrary to Californian practice) is to be closed by next March and then expensively decommissioned.

Meanwhile, Soviet *glasnost* has been followed by a wave of public concern and even protest at the widespread use of nuclear power. After Chernobyl (and with the evidence still plentiful that industrial management is a business for amateurs in the Soviet Union), alarm is understandable, but throwing stones at nuclear power has also become a polite way of complaining at the Brezhnev era. Yet the setting up of the State Environmental Commission earlier this year has now institutionalized this drum-beat of discontent. Does all this spell the end of the exploitation of uranium as a fuel? Why has the dream of the 1950s so quickly been blighted?

The British case is, modestly, rich with pointed half-answers (see page 187). The Berkeley plant was commissioned in 1962 with a design lifetime of 25 years. Simple arithmetic shows that it has done a little better than planned, if only just; the past few years must have been a plant manager's nightmare, so that many professional engineers as professional environmentalists may welcome the decision to close the plant. But what needs to be understood is that the remaining seven nuclear plants of that generation will be kept going at least until 1992 (and some will even see the new century in). Both commercially and technically, they have been a success. If the cores of the two reactors at Berkeley should not have been dismantled a century from now, they may nevertheless serve as splendid monuments to a memorable time when innocence and optimism went hand in hand.

Innocence

Innocence is endearing, but a poor guide to technological development. Even before work at Berkeley began, the then British government had issued a white paper (in 1955) outlining a ten-year programme of construction beginning with low-temperature gas-cooled reactors and finishing with the commercial fast reactors. A year later, the then Conservative government's agents (mostly scientists and engineers) were busily reorganizing Britain's creaking engineering industry, already manifestly incapable of building ships or machine-tools competitively, for that herculean task. Prototype reactors of all kinds sprang up at the nuclear establishments.

Even before the Berkeley reactors were commissioned, the nuclear establishment had settled on the design of its immediate successor — now known as the Advanced Gas-cooled Reactor (AGR) — which offered only modest technical improvements at immodest cost and whose realizations have been a millstone around the British electricity industry's neck ever since. By the early 1970s, the utilities were saying that they wanted to build

water-cooled reactors, like most other people; nearly two decades later, the first such machine is now being built in eastern England, while a public inquiry into a proposal to build a second (on the Severn estuary) will open in October. There is a sense in which three crucial decades of development have been wasted.

In Britain, the innocence of the time was not confined to the infant nuclear industry, but was endemic and industrially corrosive. The then current delusion, tinged with chauvinism, was that innate British cleverness would somehow enable a tradition of improvisation to reduce the high cost of developing effective new products in novel fields. Yet there had been ample warnings of the dangers of such short-sightedness — the world's first civil aircraft driven by jet engines (called the comet) might have been a commercial success if somebody had worried about metal fatigue. More importantly, the British steel industry (now in good shape) might have avoided two decades of trauma if half the effort devoted in the early 1960s to direct reduction (of iron ore by hydrogen) had been spent on the more accessible technique of using liquid oxygen in steel-making. So Berkeley, for all its past utility, will also be a monument to the time when people in Britain believed themselves immune from the principle that success in high technology depends on the quality of its products, which in turn depends crucially on research and development. Has that lesson yet been learned, people should be asking?

Embarrassment

The 1960s also saw the origins of a continuing embarrassment, not exclusively British, about nuclear wastes. By the end of that decade, with the pressure on public finances growing, British work on the process for incorporating highly active wastes into glass was given less attention than it deserved. The nuclear industry, harassed by the self-inflicted problems of making AGR reactors work and, no doubt, distracted by general concern that the British economy might not be sustained for very long, lapsed into casual unconcern about the difficulties.

Now, the responsible British organization NIREX, a consortium of the interested utilities, is concentrating on finding somewhere to bury its members' stock of contaminated fuel cladding and other waste classified as intermediate in level. From time to time, talk revives of burying highly-active wastes in the seabed, but the material that will eventually have to be disposed of remains in storage tanks at the reprocessing plants. This part of the nuclear fuel cycle has become the prisoner of community concern that, even if nuclear power has become a fact of life, those who use the electricity it generates need not have to put up with the inconvenience or even anxiety of having an underground repository nearby. But there are no votes in nuclear disposal, which is why politicians and even governments are willing to let these issues slide, not merely in Britain but everywhere.

The irony is that there is certain to be a mounting demand for more nuclear generating capacity a decade or so from now. Whatever the further gains to be wrung from fuel efficiency and the modest contributions to power resources from wind and solar energy, the demand for electricity will grow and must, in the foreseeable future, be met from coal or uranium. Thermo-nuclear fusion is too far off, the cost of oil is likely to be too high.



Quite apart from the possibility that the burning of fossil fuel may have to be restricted for climatic reasons, there are no conceivable circumstances in which the world's economies could safely rely on a single fuel. The lesson, in essence simple, is that it is in everybody's interest that the two clouds looming over the nuclear industry worldwide — waste disposal and safety — should be securely banished. And quickly.

The waste disposal issue is by far the simplest. Technically, it would probably be safest and simplest to tip the huge quantity of contaminated material classified as intermediate (less than 4×10^9 Bq per kg of beta activity and less than 3 times as much of alpha activity) into some deep ocean trench and forget about it. But the small risk that it might not be possible permanently to forget makes the issue, in the present state of public clamour, too sensitive to be reopened. (The British stopped ocean-dumping three years ago in the face of public pressure, with the National Union of Seamen in the van.) NIREX has done a good job in keeping the world at large informed, but cannot forcefully make the undodgeable point that, having enjoyed the benefits of nuclear power for thirty years, the British community cannot now shrug off the inconveniences that follow. The same principle applies elsewhere, notably in the United States. The disposal of the highly-active wastes, on the other hand, demands a research and development programme, preferably international.

The safety issue is the hard nut to crack. Chernobyl made the obvious point that radioactive contamination is simply a national matter. The difficulty is that the accident also reinforced a lesson learned from previous decades that the safety of nuclear plants depends more on the way they are managed (or mismanaged) than on the technicalities of their construction. How can people in one country be assured that their neighbours, even distant neighbours, run their plants properly? Making sure that there is everywhere such vociferous local opposition that local managements will tread a narrow path is one solution, but counter-productive: it may or may not engender vigilant management, but it impedes construction and even safety.

Is it too much to ask that the international agreements reached in 1986 on trans-frontier fallout might be extended to cover other aspects of safety, the quality of management included? The International Atomic Energy Agency (IAEA) at Vienna already has a kind of voluntary consultancy agency in place. Could not this device, originally intended to help developing countries happening to acquire reactors, be made, by international agreement, obligatory? Why not even go so far as international certification of competence in the management of nuclear plants? Without some device of that kind, the disjointed condition of the nuclear industry worldwide may persist until that time, next decade, when the need to build nuclear plants is recognized to be so urgent that people have to build them in a hurry — which could be really dangerous.

The parochial question of whether Britain has learned to fight shy of over-optimism is less easily dealt with. British manufacturing industry has remarkably become more effective and profitable in the past few years, but is still skimping on development, apparently indifferent to the demonstration, first in West Germany and then Japan, that making headway in the modern world requires more than mere imitation of what others have already. Fears that this truth has not yet been grasped are only reinforced by the opinion last week (see page 186) from the British government's brand-new advisory committee that Britain might win a place for itself in the exploitation of optical electronics (optical fibres, semiconductor lasers and the like) if only British industry would exert itself a little more (and if the Science and Engineering Research Council, whose real business is to support academic research, were to set up an internal organization to coordinate support). Yet two of the chief players in the field, British Telecom and the Ministry of Defence, have been sharply reducing what they spend on development in this field. On the face of things, the false conceit that difficult tasks are not really that difficult still persists. □

Keeping tabs on people

A call for revived interest in population registers will go unheeded unless abuses can be prevented.

Mr Philip Redfern, deputy-director of the British Government's Office of Population Censuses and Surveys until five years ago, is evidently the most courageous of people. Last week he took on (apparently voluntarily) the thankless task of advocating to a meeting of the Royal Statistical Society the virtues of a central register of population, not just for Britain but for everywhere. The proposal is one that should appeal to every man and woman of reason, but which most reasonable people will unreasonably reject. And for good reason.

A central population register is a device for storing centrally a list of people's names and addresses together with, ideally, an identifying number which they carry from birth to death. The only element of compulsion is that people are required to notify their place of residence when they move. There are obvious social benefits in such an arrangement. Population statistics, for example, are made more accurate, while censuses are more easily and cheaply carried out. (Denmark, according to Redfern, has already substituted its register.) Moreover, people wishing to evade taxes or other social obligations can do so less easily, but on the other hand are less likely to be deprived of their rights and benefits (the right to vote or social security payments, for example). Nobody will dispute that a register of this kind, if properly maintained, would assist enormously the efficient administration of government.

The other side of the coin, which Redfern faithfully described in his lecture last week, is also just that: a register would indeed make government more efficient. It is not an accident that the countries already possessed of population registers, largely the Nordic and Low countries of Europe, are precisely those in which governments have over the years won the reputation of being not simply democratic but open. Sweden, like the United States, has a Freedom of Information Act, but more to the point has a tiny civil service which contracts out much of its administrative work to essentially independent agencies. Elsewhere, it is much more plausible that a central register would be used for malign purposes, perhaps to tell who is living with whom, perhaps for coercing people away from chosen inclinations, perhaps even for more overtly political purposes.

Redfern acknowledged these objections last week, but did not give them their full weight. The underlying reason why reason does not carry everything before it is simply that too many genuinely democratic governments operate machinery that people do not trust sufficiently. The very least of the requirements to make a central register acceptable would be that it should be maintained by a separate independent agency; that might work in the United States, but would be a contradiction in terms in countries such as Britain, where government is constitutionally sovereign until thrown out. Yet curiously the British government just now is planning a new system of local taxes that will require local authorities to maintain just such local registers, defective (from his view) in not being coordinated, which is certain to be the cause of further ructions.

But may there not be a solution, at least in the European context, that would satisfy both the Redfern view and that of those who fear further infringements of liberty in the cause of administrative efficiency? The European Commission in Brussels has a written constitution and a proven ability to finance administrative organizations that function independently. The commission also has an interest in the free movement of people between jobs and even residences within the borders of the European communities. Would it not make sense that the commission should go into the population register business, but on the strict understanding that its register would be used for strictly civil purposes declared in advance? □

IVF research may yet receive federal funding

- US Department of Health to end moratorium
- White House opposition anticipated

Washington

THE US Department of Health and Human Services (HHS) has taken the first steps to end an eight-year *de facto* moratorium on federal funding of research on human *in vitro* fertilization (IVF). Robert Windom, Assistant Secretary for Health, last week announced that the department would reestablish its Ethics Advisory Board (EAB), a body required by department regulations to approve all research grants for IVF, but whose charter lapsed in 1980.

The decision to create an ethics board goes back to 1974, when the National Research Act established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. In 1979, the EAB produced a report concluding that it was ethically acceptable to conduct research involving human IVF and embryo transfer under certain conditions.

According to Charles McCarthy, now director of the Office for Protection from

sion for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. But without approval from an ethics board, NIH could award no grants for IVF research.

In 1980, NIH officials began a campaign to reestablish the EAB, but these efforts got nowhere within the Reagan administration. President Reagan has been extremely sympathetic to the so-called pro-life lobby, and there was little enthusiasm for research that might upset a strong political constituency. The failure to provide a new charter has proved frustrating for NIH officials. In a 1986 memorandum to the office of the assistant secretary for health, Duane Alexander, director of the National Institute of Child Health and Human Development, criticized the HHS's failure to follow its own regulations by failing to reestablish an ethics board.

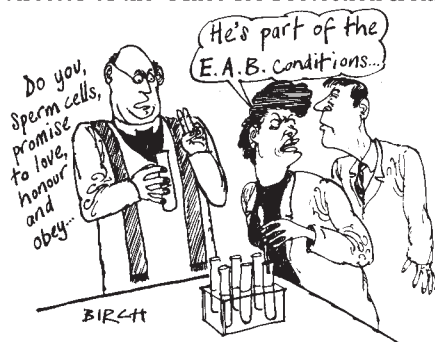
It is difficult for the HHS to maintain the respect of the research community when it insists that its regulations for protecting human subjects and animals be followed strictly by its grantees, yet ignores the parts of its own regulations that impose requirements on itself. Earlier this year, NIH tried a new tack. NIH director James Wyngaarden suggested in a memo to HHS in January that research proposals could be approved on the basis of the EAB's earlier recommendations. Receiving no reply, Wyngaarden sent a second memo in April proposing approval of a specific grant request from Washington University in St Louis to improve the culture medium used to nourish the early fertilized ovum and improve the success rate of the procedure. HHS has decided to postpone approval of the Washington University grant.

But the decision to reestablish the EAB, announced at a House of Representatives committee hearing, nonetheless came as a welcome surprise.

Aware that the decision might raise hackles at the White House, HHS took pains to maintain a low profile until the announcement was made publicly. HHS plans to publish a new charter for its Ethics Advisory Board in the *Federal Register* within 30 days, and a 60-day comment period will follow.

Choosing the board's membership will certainly take several additional months, so it will undoubtedly fall to a new White House administration to get the EAB up and running.

Joseph Palca



Research Risks at the National Institutes of Health (NIH), but formerly the staff director for the EAB, one of the conditions was that the sperm and eggs come from married couples only. In addition, the EAB concluded that IVF research that did not involve implantation of embryos could proceed so long as the research was designed to establish the safety and efficacy of embryo transfer, that the donors were informed of the nature of the research, and that no embryos were sustained *in vitro* for more than 14 days, the period normally associated with complete implantation.

The EAB report was controversial. HHS received some 13,000 comments, most of which merely stated opposition to IVF and embryo transfer research. When its charter expired in 1980, the EAB was dissolved, and its budget given to the newly formed President's Commis-

NIH's ethics panel

THE US National Institutes of Health last week announced that a panel to investigate the moral, legal and ethical aspects of using human fetal tissue from induced abortions for therapeutic purposes will meet on 14-16 September.

The panel was created at the request of Assistant Secretary for Health Robert Windom, who last March ordered a temporary moratorium on research involving the transplantation of such fetal tissue pending a report of the newly created panel. Arlin Adams, a retired judge of the US Court of Appeals will chair the panel. Dr Kenneth Ryan of Harvard University Medical School and Dr LeRoy Walters of Georgetown University will chair scientific and ethical sessions during the September meeting.

J.P.

■ The fetus and pain. See this week's Correspondence, page 190.

New genome money

THE National Institute of General Medical Sciences has awarded the first round of grants from the \$17.2 million appropriated by Congress for gene mapping. Although the NIH reckon they spend nearly \$300 million annually on research related to gene mapping and sequencing, the new awards are the first from what will ultimately be a separate project office at NIH. The largest award (\$723,160) goes to David Schlessinger of Washington University in St Louis for human genome analysis using the yeast artificial chromosome technology developed by Maynard Olson. The next largest (\$540,410) is to Frederick Blattner at the University of Wisconsin for determining the complete genome sequence of *Escherichia coli*.

The current awards represent the first year of grants that will last from three to five years. The next round of grants will be announced in September.

J.P.

Hoechst wins battle

AFTER a three-year struggle, the West German company Hoechst AG has finally overcome local opposition to its DM50-million human insulin production facility in Frankfurt. The facility will use bacteria incorporating recombinant DNA to produce the insulin, a procedure that has been used since 1982 in other countries.

A decision by the *Land* (state) of Hesse on 8 July to allow a two-year test phase of the facility included concessions to environmentalist groups which fear the release of recombinant DNA into the environment. But the chief opponents of the plant, the citizens' initiative 'Höchster Schnüffler und Maagucker,' is considering a further court appeal.

Only one other company, Boehringer Ingelheim, produces pharmaceuticals in West Germany using recombinant techniques. They make the heart drug TPA (tissue plasminogen activator).

S.D.

Dollars for biotech

BIOTECHNOLOGY has received a boost from the University of California (UC) Biotechnology Research and Education Programme. The programme, founded in 1985 and supported by state funds, has awarded \$1.3 million in annual grant money to 12 UC biotechnology research projects at seven of the university's nine campuses.

Major science research grants of \$100,000–\$350,000 were awarded for projects ranging from protein design to development of disease resistance in tomatoes. The programme also awarded smaller grants of up to \$40,000 for the investigation of issues related to biotechnology, such as a study of options for the commercialization of scientific advances.

M.B.

Optoelectronics push

THE time is ripe for a major increase in spending on research in optoelectronics in Britain — according to the government's chief advisory committee, the Advisory Council on Science and Technology (ACOST). In a report issued this month, ACOST says that "optoelectronics is vital to our future as an industrial nation" and merits more support, both public and private.

But there is unlikely to be a large injection of public funds, since the government is eager that industry should lead the way with projects determined by market criteria. The report recommends that the government should contribute to a series of demonstration projects of the commercial applications of optoelectronics.

It also recommends that the Science and Engineering Research Council should set up an optoelectronics directorate and give priority to optoelectronics in its scheme for founding Interdisciplinary University Research Centres.

C.McG.

Promising artist

LEONARDO da Vinci's genius in the fields of science and art is to be the subject of a large-scale exhibition in Britain next year. At the heart of the exhibition will be the largest range of Leonardo drawings ever assembled. Also on show will be large-scale models of his architecture from the Museum of Fine Arts in Montreal and the Science Museum in Milan. And specially designed for the exhibition is a flying machine with a wing span of 36 feet, based on Leonardo's drawings.

There will also be computer graphics sequences, made under the direction of Philip Steadman of the Open University, including sequences of Leonardo's architectural projects and a study of the problems of perspective geometry he faced while painting the Last Supper. The exhibition, sponsored by IBM UK Ltd, will be held from 26 January to 16 April 1989 at the Hayward Gallery, London.

C.McG.

JPL enthusiastic over new slant to parallel computing

Berkeley

THE Jet Propulsion Laboratory (JPL) at Pasadena is hoping that its latest parallel computer, the Mark III Hypercube, will give parallel architecture machines a lead over conventional supercomputers both in speed and in the range of problems for which they can be used.

Although parallel machines have for some time been catching up with the Crays of this world, their advantage has yet to be demonstrated.

The Mark III Hypercube is the third generation of parallel computers developed at JPL, which is owned by the National Aeronautics and Space Administration but operated by the California Institute of Technology. Like its predecessors, the new machine is based on the hypercube architecture, which connects processors as if they were at the corners of a cube in n dimensions.

When completed, the Mark III will have 128 processing nodes, with an architecture based on a 7-dimensional cube, so that each processor will be directly connected to seven others.

With a peak speed of 2,000 million floating point operations per second (2 gigaflops), the Mark III is 10 times more powerful than the NCUBE-10, formerly the most powerful parallel computer, says Carl Kukkonen, director of JPL's Center for Space Microelectronics Technology. Each processor in the Mark III is a Motorola 68020, with 16 megaflops and about four megabytes of memory.

One objective of the design of the Mark III, said Kukkonen, was to use commercially available parts to keep costs down and at the same time to follow a design that can theoretically be scaled up to create a much more powerful computer. The result, he said, is a performance that already "is equal to and slightly exceeds conventional supercomputers". Comparable in power to the most powerful Cray supercomputer, at \$2 million the Mark III is one-tenth the price.

Parallel processing computers are often limited to certain types of tasks because software to use them is limited. But Kukkonen says JPL scientists have completed 200 different applications that can be run on the hypercube, demonstrating that it is a broadly usable machine.

The next goal in supercomputers is a teraflop (1 billion million flops) machine, and the hypercube can theoretically reach that goal, says Kukkonen. "We can easily imagine going to 10,000 nodes", he says.

JPL is also seeking to develop a machine comparable with the Mark III that is reliable enough to function in space, preprocessing the telemetry from

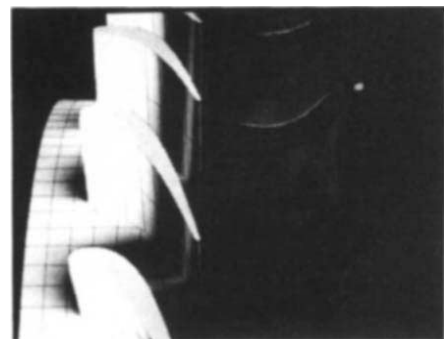
remote sensing satellites. The problem with current parallel computers is that when one node fails, the machine ceases to function. Replacing boards is not a practical matter in space. JPL is therefore developing an ultra-fast switch not only to assist the function of large parallel machines by allowing communication through nodes at 100 times the present rate, but also to allow fault tolerance — allowing messages to be routed around bad nodes and substituting spare nodes when necessary.

Marcia Barinaga

Jet simulation

Washington

COMPUTATIONAL techniques developed by Dr Man Mohan Rai of the NASA/Ames Research Center in California have made possible a three-dimensional numerical simulation, on a Cray-2 supercomputer, of gas flows past the rotating turbines of a jet engine. Rai's computer programme can simulate a pair of turbines, one stationary and one rotating, each of which is a cylindrical drum mounted with airfoil-shaped blades. In the illustration, one set of blades is moving vertically with respect to the other, and the contours show the temperature in the gas flow. The new mathematical technique that allows the fluid-dynamical



simulation to be done accurately in a reasonable length of time is a method of matching fluid properties (velocity, temperature and pressure) across the boundary between two fine-meshed grids of points, one static, the other moving with the rotating turbine. Even so, the simulation took 100 hours on the Cray-2.

Rai's work has been done with the support of engine builders such as Pratt and Whitney and General Electric, and at a press conference last week NASA officials advertised the research as an example of how federal funds can help US competitiveness. Although Rai's mathematical inventions are to be published, the computer programme itself will be treated as sensitive technology, distributed to US companies but not sent abroad without approval.

David Lindley

Closure of Berkeley nuclear plant signals decline of the Magnox

London

THE closure of Britain's first strictly commercial nuclear power station was announced last week, following a critical safety review last year by the Nuclear Installations Inspectorate (NII, see *Nature* 332, 772; 1988). The Central Electricity Generating Board (CEGB), which runs the station, has decided that the costs of implementing the safety requirements are too great.

The station, at Berkeley in Gloucestershire, was commissioned in 1962, at the same time as another at Bradwell in Essex. Both were subject to long-term safety reviews when the CEGB indicated it wished to keep them operating beyond their expected lifetime of 20–25 years. For Berkeley, the NII listed 17 key requirements that CEGB would have to

NII says it expects that work similar to that required on Berkeley and Bradwell will be necessary on the other Magnox reactors. Eddy Ryder, chief of the inspectorate, says there are common problems because of ageing and tightened safety standards. But the future of these reactors will depend on whether CEGB and the separate Scottish utilities decide it is economic to carry out the work. Exactly how decommissioning will be carried out at Berkeley, and the cost, is not certain, but something will have been learned from the dismantling of a reactor at Sellafield in Cumbria.

One reactor at Berkeley will be shut in October, the other next March. CEGB expects that half of the 532 staff will be occupied for five years in removing fuel from the reactor and its cooling ponds.



Industrial archaeology: one of the options for Berkeley is conversion into a museum.

meet in order to run the station until 1992; the cost of the work was estimated to be around £4 million.

The verdict on Bradwell was similar. Improvements are going ahead at Bradwell, where ways have been found of making them without shutting down the reactors for long periods. But at Berkeley, where a series of technical problems over the past five years has made the station's operating costs twice as high as the average of the CEGB's other seven Magnox stations, the plant will be closed.

Berkeley has been giving the CEGB problems for some time: since 1983 it has been necessary to shut down its reactors for refuelling because of faults in the main crane used to service the reactor, while one out of eight coolant gas circuits on each of its two reactors is permanently closed and gas pressure reduced by 10 per cent. Total output is thus reduced from 276 MW to 200 MW. All eight of CEGB's Magnox stations together provide half of Britain's nuclear electricity, itself 16 per cent of total generation last year.

The next safety report (on Hunterston A in Scotland) is due in the autumn, but

Thereafter, it will take 5–7 years to remove the external buildings, leaving only the sealed reactors. Stage three, dismantling the reactors, will be postponed for up to a century. The entire process is expected to cost around £300 million, compared with the £450 million set aside by CEGB. But, with privatization of the electricity supply industry imminent, it is not clear who will pay for the decommissioning of the CEGB's other seven Magnox reactors.

In what seems a bid to win over public opinion, CEGB is considering plans to turn Berkeley into a museum instead of returning it to a green-field site. CEGB says it is following with interest developments at Chinon, in France, where the same idea has been canvassed.

Christine McGourty

■ Another cloud hanging over Britain's nuclear industry is the disposal of radioactive waste. NIREX, the government's waste disposal agency, expects that 1.5 million m³ of low-level waste, such as discarded protective clothing, will be produced between now and 2030. But the one site now in use will hold only a third of

Register pleases epidemiologists

London

THE United Kingdom and Ireland are the only members of the European Community with no form of population register. Administrators and politicians in Britain have long avoided the issue, fearing public opposition. But in a report published last week, Philip Redfern, formerly of the government's Office of Population Censuses and Surveys, says they may not be able to do so much longer. With the removal of frontier controls and the free movement of people and trade, exchanges of personal data will increase and greater compatibility between national record systems will be necessary.

Redfern argues the benefits a register would bring to society, the government and to statisticians. And it would not affect people's privacy, he says, as a wide range of personal data is already held by public authorities.

But a lack of coordination between them leads to duplication and inconsistency. A centralized register would bring some order to this ramshackle system, says Redfern. It would put a brake on fraud, crime and illegal immigration and lead to a fairer society where duties were fairly shared and benefits and rights go to those entitled to them. A register would save the government money, he says, by reducing administrative costs and cutting losses from tax evasion and improperly paid benefits. Implementation of policies would also be easier: data would already exist on which to base a cancer screening programme, for example.

Christine McGourty

that. And for intermediate-level waste, such as irradiated fuel cladding and reactor components, NIREX has not yet found an acceptable means of disposal. By 2030, 250,000 m³ will have accumulated. At present it is stored where it is created. High-level waste is similarly in storage, pending a decision on disposal.

Last year NIREX circulated a discussion document outlining three possible options for disposal of low- and intermediate-level waste: a deep-mined cavity below land or an offshore cavity below the sea-bed reached either by a tunnel from the coast or by a drilling platform. Site proposals are expected later this year.

But plans for a deep disposal facility have already been set back; the local authority has refused planning permission to British Nuclear Fuels Ltd, which wanted to drill at Sellafield in a preliminary geological investigation. NIREX abandoned plans for a shallow dump over a year ago because of the strength of public opposition.

C.McG.

Ridding space of flotsam

Washington

AN exhibition of spacecraft and satellite parts damaged by orbiting debris, including a section of window from the space shuttle Challenger found to have been cratered by a chip of paint, was the centre-piece of a Congressional hearing last week.

The House of Representatives committee on Space Science and Application was learning about the hazards to spaceflight posed by the accumulated detritus of human activity in space. According to the National Aeronautics and Space Administration (NASA), a vehicle orbiting at 300 km altitude can expect to suffer a damaging collision once in 71 years, but if the density of debris increases at the present rate, that time will fall to 8 years in 2010.

The Department of Defense has been tracking this space debris for some time, and estimates that there are now more than 7,000 objects bigger than 10 cm in low Earth orbit. But the greater danger is from smaller pieces, which are harder to detect and much more numerous. At velocities of up to 15 km per sec, microscopic fragments can cause fatal accidents.

Although low-flying debris gradually falls to Earth through atmospheric friction (a process quickened in the past few years as increased activity associated with the approaching solar maximum has puffed up the outer layers of the atmosphere) little can be done to clear away the mess.

The chairman of last week's subcommittee hearing, Representative Bill Nelson (Democrat, Florida), was concerned to find ways to slow the accumulation of debris, especially as more countries and independent organizations begin to launch commercial satellites.

The Federal Communications Commission (FCC) regulates satellite launching in the United States, but at present has no power to demand changes in design that would reduce 'littering'. Courtney Stadd of the Department of Transport speculated that FCC might be given greater control not only over the orbital placement of new satellites but also over their construction, even to the extent of disallowing painted signs on their surfaces.

Without international agreement, moves by the United States alone will not solve the problem, but Michael Michaud of the State Department's Office of Advanced Technology argued that US action against space debris would provide a base for international action. Although no immediate legislation is likely, there is a feeling that long-term questions of US space policy, neglected in the effort to get the shuttle relaunched, will soon demand answers.

David Lindley

Endangered squirrel threatens Arizona observatory

Tucson

To defend an endangered squirrel, the US Fish and Wildlife Service last Friday turned down a proposal by the University of Arizona (UA) to build seven telescopes on top of an 11,000-foot peak in southern Arizona. But a smaller observatory might be allowed.

For eight years, UA's Steward Observatory has eyed the peaks of Mount Graham for a major new observatory that would include the 11.3-m binocular Columbus telescope and the 9-m submillimetre telescope funded in part by the Max Planck Institute. Peter A. Strittmatter, director of Steward Observatory, says the mountain-top location is desirable because it has good viewing conditions, is close to Tucson and has a road for easy access.

But Mount Graham is also a unique desert 'sky island', an outpost of old growth spruce-fir forest with a smattering of flora and fauna from Mexico, said Randall A. Smith, forest wildlife biologist with the Coronado National Forest. It has the densest population of black bear in the southwest and is the only home of the Mount Graham red squirrel, *Tamias-*

ciurus hudsonicus grahamensis, classified as endangered in June 1987. That squirrel became the focus of a biological assessment by the U.S. Forest Service and of the Fish and Wildlife Service. Since 1984, UA has juggled telescope numbers, acreage and peaks to gain access to Mount Graham. Originally it asked for 40 to 60 acres for 13 telescopes spread over five sites, but in late 1986 asked for 10 telescopes on two peaks, High and Emerald Peaks. After assessing the biological impact, the US Forest Service suggested the observatory be limited to five telescopes on High Peak alone. UA responded that it had to have seven telescopes to make the observatory worthwhile.

According to the Fish & Wildlife Service, those seven telescopes "will jeopardize the continued existence of this endangered species" and therefore would violate the Endangered Species Act. The opinion contained three "reasonable and prudent alternatives" that would allow some telescopes and still protect the squirrel. The opinion is a legally binding document that limits the options to those given in the opinion, said Smith. The university must decide by 22 July whether or not it wishes to pursue any of these alternatives.

One would be to build the telescopes elsewhere; the opinion notes that, of the telescopes specified by UA, only the Columbus Project needs be built in the continental United States and that there are two possible sites in New Mexico. But UA wants the observatory in its backyard, where new technology is being developed.

The other two options allow telescopes to be built on just one proposed site, but would close the mountain-top to the public and force out summer residents and a religious camp nearby. Expansion might be permitted in 10 years if no harm had come to the squirrel.

UA says a useful observatory on Mt Graham must have at least 7 telescopes. It could then also spread its estimated \$11 million annual operating costs between the various partners in the telescope projects. Another option allows the university to build an unspecified number of telescopes on 14 acres on High Peak. But in such a limited space, seven telescopes would change the local microclimate, negating the advantages of the site.

Once the university decides how to proceed, telescope construction may still be prevented by the Forest Service, which must not only protect the squirrel but also balance the needs of other Mount Graham users against the special interests of the astronomers.

Elizabeth Pennisi

A small step for SERC

London

Britain has taken a step forward in space research with the approval by the Science and Engineering Research Council (SERC) of participation in three new missions over the next eight years. The council has approved expenditure of £16.3 million on two missions which form the cornerstone of the European Space Agency's Horizon 2000 programme, dedicated to the study of solar terrestrial physics. These are the SOHO and Cluster missions. SOHO is due for launch in 1995 and will study the outer region of the Sun's atmosphere and the solar interior. Cluster, due to begin in late 1995, will send four spacecraft into orbit round the Earth to study the evolution and properties of the plasma structure of the Earth's atmosphere.

SERC also approved involvement in a collaborative project with the Soviet Union. The Soviet Spectrum-X satellite, to be launched in 1992, will carry X-ray instruments for British spectroscopy experiments, with SERC spending £6 million on the mission. The third project approved involves participation in the Japanese Solar A mission. The satellite, due for launch in 1991, will host instruments for the study of high-energy phenomena in solar flares, concentrating on the impulsive phase when the prime energy release occurs.

Christine McGourty

Battle to save education enters critical phase in north Germany

Düsseldorf

THE battle to maintain a vital educational system in the face of acute economic pressure is being waged with particular vigour in the northern German *Land* of North Rhine-Westphalia (NRW). The government is moving into the critical phase of a 15-year plan — the 'university structure plan 2001' — to reorientate its universities toward high technology.

Elections approaching in 1990 on both state and federal levels have spurred North Rhine-Westphalia's Social Democratic (SPD) government to put its programme in place by next year. The task seems next to impossible: high unemployment, huge budget deficits, the expense of new equipment and the political resistance to eliminating courses of study pose enormous problems.

Northern West Germany and NRW in particular have suffered under the rationalization of the steel and coal industries, and new technology-oriented companies have primarily settled in southern West Germany in the years of rebuilding following the Second World War.

Science Minister Anke Brunn (SPD) has dealt with the problem pragmatically since inheriting her post in 1985. The overall strategy since the early 1980s has been to cut unnecessary courses in teaching or humanities, where students would not find jobs anyway; to create regional centres in both natural sciences and humanities and reduce duplication in neighbouring universities; and to increase numbers of graduates in fields where there are expected to be plenty of jobs, such as electronics or information sciences.

To the great relief of the universities, which are suffering from student overpopulation as a result of the post-war baby boom, Brunn has warded off the Finance Minister's budget axe and ensured that personnel will not be cut until at least 1990. Aware of the risk that frustrated students might go to universities in southern Germany (which is enjoying something of an economic boom), Brunn is attempting to make NRW's universities as attractive as possible through a publicity campaign.

Last month, university administrators were bowled over when they found out that their budgets for small- and medium-sized equipment would be increased dramatically after a freeze that had lasted several years. NRW had to go further into debt to finance the equipment renewal programme, another sign that it considers research a "new raw material" for the future, as NRW Minister-President Johannes Rau (SPD) recently announced.

But the reorganization plans have met with resistance from university rectors who feel that they are losing their autonomy to the Science Ministry. Although they agree that reorientation is necessary, the rectors reject the Ministry's heavy-handed, insensitive approach. Rector Klaus Habetha of the Technische Hochschule at Aachen regrets that Brunn has not called for expert advice from the universities themselves in deciding where to make the next round of cuts, which will involve a broad range of fields especially



Anke Brunn — pragmatic approach

in the humanities and social sciences. In future, he said he hopes that cuts will not just be "decreed from the top" but rather that "a common solution" can be achieved.

Rector Knut Ipsen of the University of Bochum complained that a small university such as Bochum, which offers a lot of small "exotic" subjects in humanities and the arts, suffers disproportionately under across-the-board cuts. For example, under the Ministry's plans, the department of modern Greek in Bochum would be reduced to one half position. He considers such cuts "an absurdity".

Brunn put the situation in a much more positive light. In the natural sciences and engineering, NRW has "no problems" attracting and holding onto good researchers. The expansion of the 1970s created high standards that have been maintained, she said. The situation could be improved still further, she says, by a fairer distribution of federal research funds. In the ongoing struggle between the federal and *Länder* governments for funds Brunn feels that NRW has been shortchanged. She attributes this to regional bias on the part of referees.

Though the situation is improving as more NRW researchers sit on grant committees, Brunn still feels that NRW is discriminated against by the Max Planck

Sagdeev to leave Moscow post

London

ACADEMICIAN Roald Sagdeev, director of the Space Research Institute in Moscow, is planning to step down from his post in October, after an election for a successor conducted in the true spirit of *perestroika*.

This development is important both because it is meant, in part, as an example to others within the Soviet Academy of Sciences system and because Sagdeev has been influential in fostering relations between Soviet science and the rest of the scientific community during the past four years. Sagdeev said this week that he plans to become a working scientist after having been at the Space Research Institute for fifteen years.

Under the rules governing the appointment of the directors of academy institutes, institutes can only recommend the successor of a retiring director to the praesidium of the academy, which retains the power of preferment. But in the circumstances (and the present climate) it is hard to think that the recommendation would be declined.

Sagdeev's intended departure will nevertheless be a disappointment for many scientists in the West who have benefitted from his enthusiasm for carrying their instruments on Soviet spacecraft.

Among Soviet research establishments, the Space Research Institute in Moscow stands out for its efficiency and effectiveness. Sagdeev himself has been personally responsible for much of the scientific success of the Soviet space programme in the past decade and more.

Even by present standards, Sagdeev is something of a radical. He was elected a delegate to last month's special conference of the Soviet Communist Party, at which he proposed the amendment to the official roster of resolutions requiring that the ten-year retirement rule should apply even to the general-secretary of the party (sitting over his left shoulder), with whom he is well-connected.

The chances are that Sagdeev's intention to step down from a post he has made influential will not inhibit his frequent appearances in the West or even diminish his personal influence on the course of East-West events. Latterly, he has been contributing regularly to the *glasnost* newspaper *Moscow News*, which may be a sign of what lies ahead. **John Maddox**

Society. In the humanities and social sciences, the crisis mentality pervading the universities cannot be attributed to the Ministry, said Brunn. According to her statistics, there will be over 2 million academics in West Germany by 2000 but only 800,000 positions.

Steven Dickman

Recognizing good teaching

SIR—In the 17 March 1988 issue, *Nature* reviewed the achievements of secondary school students in biology, chemistry and physics in 17 countries¹, based on a US National Science Foundation (NSF) survey². The ranking is startling, and you quote appropriately that “for a technologically advanced country [USA], it would appear that a reexamination of how science is presented and studied is required”³. Improving education is an important and complex task.

A recent issue of the *Chronicle of Higher Education*³ gives a 30-point guideline for the measurement of excellence in a teacher. Only the eighteenth rank is assigned to the criterion “are knowledgeable about their work”. I believe mastery of the subject is the most important requisite. The article quotes a professor as saying that teacher evaluation “emphasizes the superficial”. Among the 30 points, nothing is said about course content relative to students’ preparation and need. The guidelines do not expect an excellent teacher to teach principles rather than data. The good teacher does not teach by rote but must show the ways hypotheses are conceived, tested and applied. In the NSF report, it is not considered a very important feature of a good teacher that he or she weighs the subject, integrates the facts into theory and is up-to-date and relevant. Is it not an absolute necessity to use plain and expressive language, rather than to “give corrective feedback promptly to students”? It is less important that the teacher should “provide clear and substantial evidence that the students have learned” than that the students feel they have gained by the instruction. The measure of the effectiveness of teaching is not the enthusiasm of the instructor but how much interest and enthusiasm he can generate. The good teacher inspires the students, is patient with the learner, shows interest in individuals and cares for them.

Teacher evaluation is important and cannot be dealt with in the way suggested by the guidelines. The *Chronicle* aptly quotes someone who says that one such evaluation “generated mostly paperwork”. It is bad enough that at many institutions research is judged by number of publications rather than by their quality and/or impact. Some administrators advocate evaluation based on “market value”, thereby admitting their inability to make a sound judgement, and confusing commodities such as eggs with eggheads. There is one solid criterion of an excellent teacher: her/his students know more than those of an average good instructor and have learned by less painful ways. They understand the subject better and can apply the principles more effectively and

ethically. This gain by the students can be objectively assessed, and the best teachers can thereby be recognized.

G.P. REDEI

*University of Missouri,
Columbia, Missouri 65211, USA*

1. Ezzell, C. *Nature* **332**, 195 (1988).
2. Science and Achievement in Seventeen Countries. (US National Science Foundation, Office of Studies and Program Assessment, Washington, DC 1988.)
3. *The Chronicle of Higher Education* **34**, A13, A15 (1988).

Opportunities in India

SIR—K. S. Jayaraman’s article (*Nature* **333**, 201; 1988) about an apparent halt of the brain drain from India distorts the facts. It is time that these ambiguities were clarified to avoid any further wishful thinking on the part of the scientific community in general and the Indian Department of Science and Technology in particular. The fact is that the apparent halt in the brain drain is a consequence of changes in Western immigration policies. In the mid-1970s, immigration laws in the United States and United Kingdom were lenient, allowing graduates to move to these countries. Immigration laws are now much more stringent; only 2 per cent of Indian graduates are taken by these countries. It is therefore wrong to assume that the decrease in the brain drain is due to an increase in opportunities in India.

Perhaps some young scientists of Indian origin who have made names for themselves would like to return to help bring science in their home country to the forefront and some of them even leave well-paid jobs to do so, but unfortunately this is at present only a one-sided effort. The Department of Science and Technology is keeping a register of graduates who wish to return to India, but the real issue is whether anything is being done to give them incentives.

Foreign graduates face two major problems. First, their way of tackling scientific research is different from that of graduates trained by the bureaucratic older generation of Indian scientists, who feel threatened by the foreign-trained invaders. Second, they have to face the ingrained bureaucratic attitudes of the older generation. These problems will not be solved unless the foreign-trained and capable young scientists have senior positions. There is a formidable number of new institutes being built by Indian science and technology, and financial support from WHO, UNO and the United States is going into the system. Real scientific research will happen only when the science in these buildings is carried out with purposeful objectives.

BALBIR S. BHOGAL

*12901 Cottage Mill Place,
Midlothian, Virginia 23113, USA*

Pain and the fetus

SIR—The debate generated by David Alton’s attempt to amend the 1967 Abortion Act ranged widely. But one question that it failed to address, which to me, as a neuroscientist, is perhaps the most compelling, is the possibility that severe pain is experienced by the fetus undergoing late termination. I have heard no informed argument on this matter and, as a result, remain worried by the possibility that as a society we are sanctioning an inhuman practice.

There is detailed legislation relating to pain and laboratory animals, and I wonder whether there is some inconsistency between that regulation and what appears to be a lack of information and controls over dealings with human fetuses.

D.E. MARPLE-HORVAT

*Department of Physiology,
University of Bristol Medical School,
University Walk, Bristol BS8 1TD, UK*

Waiting refusniks

SIR—Your leading article (*Nature* **333**, 483; 1988) asserts that those who champion the cause of Soviet refusniks may be flogging last decade’s horse.

I agree; but it is not the fault of those who try to help refusniks that the Soviet Union has failed to drag this particular horse into the 1980s. We welcome many of the changes now taking place in the Soviet Union. But we would welcome them still more if they led to the release of refusniks such as Oskar Mendelev or Vladimir Raiz, who have been waiting for exit visas since the early 1970s. So long as the Soviet Union tries to have us believe that a protein crystallographer who left his work in 1973 still retains secret information of importance to the security of the state, there will be many in the West who remain sceptical of its intentions about more fundamental matters.

M. D. YUDKIN

*Scientists for the Release
of Soviet Refusniks,
4A New College Parade,
London NW3 5EP, UK*

No energy

SIR—In “Explosive fragments by numbers” (*Nature* **332**, 775; 1988), John Maddox fails to include the translational kinetic energy of the shell. This means that, in the limiting case of one fragment (no explosion), the available destructive energy is zero. The implication would be that we have nothing to fear from a rifle bullet. Kinetic energy SDI systems would also be contraindicated.

EDWARD DERRINGH

*Physics Department,
Wentworth Institute of Technology,
Boston, Massachusetts 02115, USA*

Does the past spell out the future?

Current anxieties about the prospect of climatic change, as a consequence of the carbon dioxide greenhouse or other means, raise important questions about the limitations of statistical inference.

CAN a journal with the title *Climatic change* fairly claim to be objective? That quizzical question, born of the suspicion that the title is meant to suggest the climate must be changing when it may not be, should be struck down. For one thing, journals can call themselves by the titles with which they are most comfortable, and should ordinarily choose titles that seem to suit them best. For another, there is a continuing and serious interest in the prospect that the present climate may be unstable, first stimulated by the recognition that the coldest spell in the most recent glaciation was merely 18,000 years ago, when the English Channel was dry land.

It is entirely proper that people impressed by that and other observations of the fragility of our climate should contribute articles to a journal entitled *Climatic change*. It is also natural, in the circumstances, that many (but by no means all) of the contributors should be concerned with the possible effects of carbon dioxide on the terrestrial climate. But the journal also raises implicitly (and sometimes explicitly) the interesting question of the degree to which it is permissible to predict what the future holds for climate on the basis of observations of the past.

The occasion for this introspection is an article by R.P. Kane from Brazil's Institute of Space Research, now under the wing of the ministry of science and technology, with the arresting title "Are droughts predictable". The essence of the argument is that they are predictable in some parts of the world, but not everywhere. Starting with records of precipitation going back for decades, the authors conclude that there are, indeed, some parts of the world where dry spells have recurred at regular intervals. These include north-east Africa and north-east Brazil.

The criterion for deciding when future predictions of drought should be taken as reliable is that past fluctuations of rainfall should stand out from statistical noise by at least two standard deviations, perhaps even three. Moreover, it should not be thought that short-term fluctuations (with periods of two or three years) are reliable indicators, given their probable connection with what is called the quasi-biennial oscillation (linked, among other things, with the El Niño phenomenon). But, with these reservations, the article does go on to support its author's decision to declare where and when there will be serious

droughts by means of brackets beneath graphs which are time-series of recorded rainfall at different places.

The article is really quite disarming. Looking back, the authors note that when, in 1978, C. Giradi and L. Teixeira from Sao Paulo published a prediction that there would be a drought in north-east Brazil between 1978 and 1983, "Giradi gave interviews to the press, which gave due publicity to his views", that the "government ... and the meteorological community ignored his views", but that, "nevertheless, the drought occurred...". Plainly, there is no future in being a prophet, especially in one's own country. Kane and Trivedi go on to say that they have no wish to "scare people as by crying wolf" and that they expect that the precautions taken by people heeding their warnings will be harmless in the sense that if the predicted droughts arrive, "the precautionary measures would come in handy", but that if the droughts do not arrive, "no harm would occur".

To give the show away, the Kane and Trivedi article is exactly two years old (*Climatic change* 8, 209; 1986). That is too soon to know whether the particular predictions have been sustained — many of the droughts predicted will not arrive until the middle-1990s. But the philosophical issue raised by the article is even more important now than then: when can it be both prudent and useful to make predictions of natural phenomena? Droughts and earthquakes are obvious examples, but current interest in the greenhouse effect expected to follow from the accumulation of atmospheric carbon dioxide is the most obvious illustration.

Perceptively, *Climatic change*, recognizing the solemnity with which predictions must be regarded, invited A. Barrie Pittock from the CSIRO Division of Atmospheric Research near Melbourne, and a member of its editorial board, to comment on what it had done. Pittock, prominent among those who a few years ago were warning people of the hazards of the nuclear winter that may follow all-out nuclear war, makes an exceedingly intelligent and interesting comment on the drought predictions from Brazil, a revealing comment about nuclear winter, an ambiguous but implicit comment on proper attitudes towards the greenhouse effect and a helpful comment on the problem of statistical inference in general. Too little attention has been paid to the last of these

in recent discussions of the last-but-one.

First, there is the general principle that statistical analysis can only test the validity of hypotheses, not prove them to be correct. Moreover, there is always a confidence limit involved, explicitly or otherwise: is 5 per cent (95 per cent) sufficient, is it necessary to go to 1 per cent (99 per cent) or is the expected degree of confidence a measure of the subjective level at which one's data are likely to be significant? Pittock rightly declares that arguments such as these melt away only when the problem of inferring the truth about the real world is simplified by the arrival of a physical mechanism connecting the future with the past. If we did not know about the Solar System, we should be a little at a loss to know what to make of the diurnal rising of the sun, millennia of observations notwithstanding.

Statisticians have long been familiar with an even more teasing problem in statistical inference, that of accounting for the different behaviour of people when confronted with the same statistical evidence. Bad news about cigarette smoking has persuaded many people to give up the habit, but others apparently believe the inferences that might be drawn do not apply to them. Pittock puts the point differently by noting that on a prediction of a 40 per cent probability of rain, some people will carry umbrellas and others not, and then by remarking that the distinction is not simply one between pessimists and optimists, but that the circumstances matter; if one is off to swim, one can afford to be an optimist about the rain.

How does this apply to climate prediction? Pittock says that science "acts to protect the correctness of what is considered as established scientific knowledge" and thus rejects hypotheses until proven. But in the real world, he says, it may be "socially responsible to urge action on the basis of a hypothesis ... not proven beyond reasonable doubt". That is where nuclear winter comes in. (Kane and Trivedi come poorly out of this, for acting on false warnings could cause "a lot of unnecessary expense and suffering".)

The most worrying aspect of this is that it is a licence for allowing undeclared and subjective calculations of consequences to influence pronouncements. In the argument about the greenhouse that lies ahead, it may be prudent to ask the experts who testify on both sides what their hidden calculations are. **John Maddox**

Regulation of transcription

Flexible interpretation

N.J. Short

IT has been known for some time that a eukaryotic cell can generate alternative messages from one gene with a facility which would be the envy of any Pentagon official interpreting an arms treaty. A wide range of devices, such as alternative promoters, terminators and splicing sites, are available for this purpose, and many genes are known that use them. In a striking paper on page 218 of this issue¹, Santoro *et al.* show for the first time that the gene encoding a transcription factor falls into this category, and so uncover a new mechanism by which the effects of factor-binding sites in different cells can be varied.

In a previous News and Views article², I described the purification of a factor called CTF, which stimulates transcription from various viral and mammalian promoters, and is now known to mediate at least some of the increased transcription caused by transforming growth factor- β . The site to which this factor binds includes the sequence AGCCAA, and in its optimum form has a 2-fold axis of symmetry. Surprisingly, it became apparent during purification that the same protein is also responsible for stimulating the replication of adenovirus DNA, in which guise it had previously been named NF-1 (ref. 3). Although a purified preparation from HeLa cells was shown to act both at the adenovirus origin of replication and at several promoters, repeated rounds of affinity chromatography failed to produce a preparation with a single molecular mass, and it was concluded that the protein is inherently heterogeneous.

Family of factors

One reason for this heterogeneity is now clear. In the work described in this issue¹, Tjian and his colleagues sequenced peptides from their purified CTF/NF-1, synthesized oligonucleotide probes encoding some of the peptide sequences, and used the probes to isolate several complementary (c)DNA clones derived from the mRNA for their protein. Their remarkable conclusion is that the cDNA clones fall into three groups, which have many regions in common, but differ in two limited areas. As each clone hybridizes to the same piece of chromosomal DNA under stringent conditions, this indicates that transcripts from the one gene encoding CTF/NF-1 are spliced in at least three ways to create a family of factors influencing transcription and replication.

What is the function of the regions that are spliced in or out in different transcripts? One possibility is that they mediate differ-

ent functions, such as DNA binding, stimulation of transcription, or stimulation of replication. Tjian and his colleagues show that this is not the case by expressing their cDNA clones in *Escherichia coli* as fusion proteins. It is virtually impossible to distinguish any of the three cloned forms of CTF/NF-1 from the original purified preparation by DNase I footprinting, and all three forms stimulate the initiation of adenovirus replication. Two of the cDNA clones have also been expressed in *Drosophila* (which lacks an endogenous CTF/NF-1) and shown to stimulate transcription from exogenous promoters carrying a suitable binding site.

The functional differences between the three forms of CTF/NF-1 thus remain obscure. Nevertheless, it is conceivable that each form could have subtly different effects on transcription or DNA replication, perhaps by interacting in different ways with some of the other protein factors involved in these processes. This would then provide an organism with an economical way in which to alter the effects of a given factor-binding site in different tissues. Tantalizing hints suggesting an elaborate version of this general scheme already exist. Not only do the preparations of CTF/NF-1 from HeLa cells seem to contain more than three sizes of polypeptide³, but there are indications that transcripts encoding this factor can be spliced in different ways in different cell lines¹. Moreover, 3T3 cells and porcine hepatocytes contain only a single polypeptide capable of binding CTF/NF-1 sites with high affinity.

The situation is made even more complex by the number of proteins which can stimulate transcription by binding to sequences containing the motif CCAAT. These factors can be distinguished by such features as their relative preferences for different sites containing the motif, the DNA residues with which they make close contact when they bind, and their thermostability. There have been so many reports of such proteins, and so few comparisons between the proteins are available, that it is hard to say how many are currently under investigation. A conservative estimate would be at least five. Given that each of these factors may be made in various forms, as CTF/NF-1 is, the potential complexity of the system is staggering.

In several instances, purification of these factors has revealed that what at first appeared to be a single polypeptide is in fact a heterodimer, both parts of which are required for efficient binding of DNA. In the cases so far examined, only one of the

subunits is actually responsible for binding to the DNA, but binding is only weak without the second subunit, which forms a complex with the first and strengthens the DNA-protein interaction. The two subunits are therefore easily separated by ion-exchange chromatography.

Apparent conflict

In a recent paper⁴, Chodosh *et al.* use the gel electrophoresis retardation assay, or 'move-up', to characterize three factors of this kind from HeLa cells, one of which is indistinguishable from CTF/NF-1 by the methods they used. Their assertion that this protein bound only as a dimer was based on the need to mix two fractions prepared by ion-exchange chromatography of a whole-cell extract before any significant binding to a CTF/NF-1 site could be seen. This result at first sight seems to conflict with the observations of Santoro *et al.* in this issue that all three of their CTF/NF-1 polypeptides bind strongly to DNA and can stimulate transcription or DNA replication without further assistance. Moreover, another group has purified fully active CTF/NF-1 from HeLa cell extracts, beginning with a single fraction from an ion-exchange column⁵. This seeming conflict, however, is not as serious as at first sight. The two laboratories^{1,4} prepared their samples from different subcellular fractions by quite different procedures, and assayed them using different techniques. The method used in any of these steps can have a strong effect on the observed properties of individual factors. It is entirely conceivable that, under the conditions of Chodosh *et al.*⁴, the same protein cloned by Tjian and his colleagues¹ required an additional subunit if it were to bind efficiently.

One of the other 'CCAAT-box factors' characterized by Chodosh *et al.* binds only to a limited region around this box, and is even more dependent on a second subunit for efficient binding than CTF/NF-1 seems to be. Both these features are also characteristic of a transcription factor from yeast which is composed of two subunits called HAP2 and HAP3. Struck by this resemblance, Chodosh *et al.* compared the two factors in detail, and demonstrated that they have the same relative preferences for different forms of their binding site and make identical contacts with the bound DNA⁶. So strong did the resemblance seem that they then took the bold step of attempting to form yeast-mammalian dimers, by mixing fractions containing CP1A with HAP2, and fractions containing CP1B with HAP3. To their surprise, both dimers not only form but also bind DNA with the same specificity as the original proteins.

Although such a detailed resemblance between the transcription systems of yeast and mammals is remarkable, it is by no means unique. In an earlier News and

Views article⁷, I described the close relationship between the proteins GCN4 and AP-1 (otherwise termed c-JUN), which stimulate transcription in yeast and mammals, respectively. Although it has since been shown that, when the AP-1 binding site is introduced into yeast, it is bound by a different protein from AP-1, unmodified v-JUN has also been shown to stimulate transcription from suitable promoters if it is expressed in yeast⁸. An even more striking case is that of the yeast protein GAL4, which is capable of stimulating transcription both in mammalian cell lines and in intact *Drosophila*. This protein is particularly interesting, as it was the first transcription factor in which the protein domain responsible for influencing transcription was identified. It has since been demonstrated that only a single part of this domain is necessary and sufficient for the stimulation of transcription⁹. This is the structure which Paul Sigler, in a recent News and Views article¹⁰, memorably termed a 'negative noodle'; similar structures have since been identified in GCN4 and, although it was at first overlooked, in JUN. It is thus of great interest that CP1 is so closely related to HAP2/3, and the genes encoding CP1 will doubtless be closely scrutinized for negative noodles once they have been isolated.

CP1 is not the only CCAAT-box binding protein to resemble a yeast transcription factor. Santoro *et al.* note¹ that all three of their CTF/NF-1 clones encode proteins with amino termini fairly similar to the carboxy-terminal portion of a factor called RAP1 (also termed GRF1). RAP1 is a protein with remarkable properties first identified as binding to a so-called silencer sequence at the HMR locus in yeast¹¹. Such sequences are the inverse of enhancers, and repress transcription in both directions over long distances. It was therefore quite unexpected to find that the same protein acts in different surroundings to activate transcription, and that it can also influence the transmission of plasmids at mitosis, as well as binding to yeast telomeres^{11,12}. Multiple functions of this kind seem to be fairly common in yeast DNA-binding proteins.

Thus another factor originally identified as binding to the HMR silencer seems elsewhere to have a role in chromosome replication and in delimiting transcription units^{11,12}. Similarly, a factor regrettably also termed CP1, which acti-

vates transcription at the GAL2 locus and may well be related to a mammalian protein termed either USF or MLTF, also binds to sequences at yeast centromeres¹³.

As the negative noodle is the only structure which has so far been shown to stimulate transcription in yeast, these results suggest that it may be capable of mediating a number of other functions as well. Is it possible that CTF/NF-1 has a similarly varied set of roles in mammalian cells?

Are the stimulation of transcription, the silencing of transcription, and the initiation of DNA replication all implemented in eukaryotes by proteins which spend their time, like characters from the film *Tamopo*, searching for the perfect noodle? It will be fascinating to find out. □

N.J. Short is in the Department of Biophysics, Cell and Molecular Biology at King's College London, 26-29 Drury Lane, London WC2B 5RL, UK.

Oxide superconductors

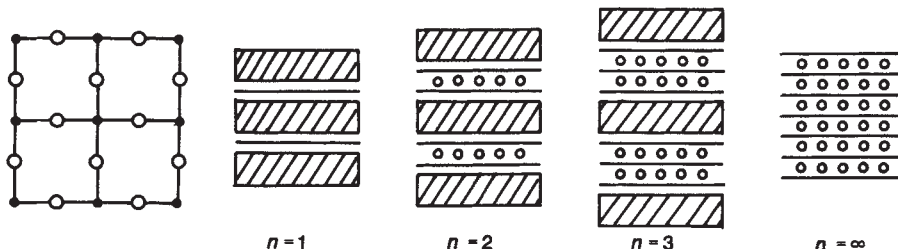
Infinite stacks of copper oxide

Colin Greaves

It is now known that $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$, discussed in an earlier News and Views article¹, belongs to a family of structurally related high-temperature superconductors which contain the elements Bi–Sr–Ca–Cu–O or Tl–Ba–Ca–Cu–O and have the general formula $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ (A=Bi, Tl; B=Sr, Ba). Structural studies of phases with $n=1-3$ have established that copper-containing blocks, each with n $[\text{CuO}_2]_\infty$ layers interleaved with $n-1$ calcium layers, are separated by

this family explains this valence restriction and reveals additional features that are unique to the $n=\infty$ member.

The structural study by Siegrist *et al.*⁵ clearly demonstrates the very simple $n=\infty$ structure (see figure) in which only calcium or strontium ions separate $[\text{CuO}_2]_\infty$ layers. This highly defective structure, a perovskite with a third of the oxygen positions vacant, has copper bonded to only four co-planar oxygen ions. The most pronounced difference



The family of superconductors $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ consists of n layers of $[\text{CuO}_2]_\infty$ (left, from above; right, lines, from the side) separated by blocks of $[\text{ABO}_2]_\infty$ (hatching). The transition temperature increases for $n=1$ to 3, but the $n=\infty$ member is insulating. ●, copper; ○, oxygen; ◦, Ca^{2+} ions.

$[\text{ABO}_2]_\infty$ units with NaCl structure (see figure and, for example, refs 2 and 3). It is notable that the transition temperature T_c increases with n ; for the thallium-based series², $T_c=80, 110$ and 125 K for $n=1, 2$ and 3 . This is compatible with one theoretical model⁴ that involves the hopping of valence-bond electron pairs between $[\text{CuO}_2]_\infty$ planes and predicts a fairly rapid decline in the stepwise increase in T_c . Nevertheless, the preparation of members with larger n appears essential for further progress and the report by Siegrist *et al.* on page 231 of this issue⁵ of the synthesis and structural characterization of $\text{Ca}_{0.86}\text{Sr}_{0.14}\text{CuO}$, represents the impressive leap to the $n=\infty$ end member of this series.

This parent material is insulating⁵ which superficially might seem to contradict the tendency for T_c to increase with n noted above. But because the copper ions are strictly divalent in this material, mechanisms of superconductivity involving mixed valence cannot apply. Close inspection of the structural chemistry of

between the parent compound and its bismuth- and thallium-containing relatives, $n=1-3$, is the absence of $[\text{ABO}_2]_\infty$ connecting units, which is likely to modify the electrical properties by the consequent changes to the copper environment and, probably, also its valence.

Siegrist *et al.*⁵ attribute the insulating behaviour to the copper valence. Although structural data for the bismuth and thallium phases imply the presence of only Cu^{2+} ions, chemical analysis of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ favours a $\text{Cu}^{2+}/\text{Cu}^{3+}$ mixture⁶. Because some Cu^{3+} appears essential for superconductivity, it has been proposed^{7,8} that Cu^{3+} either compensates for incomplete occupancy of the A or B sites, or results from electron transfer from copper bands to thallium 6s or bismuth 6p bands; the latter possibility would not explain the chemical analysis. In $\text{Ca}_{0.86}\text{Sr}_{0.14}\text{CuO}$, the absence of $[\text{ABO}_2]_\infty$ layers eliminates both mechanisms, and for this strictly Cu^{2+} compound, electron correlation effects are expected

1. Santoro, C. *et al.* *Nature* **334**, 218–224 (1988).
2. Short, N.J. *Nature* **326**, 740–741 (1987).
3. Jones, K.A. *et al.* *Cell* **48**, 79–89 (1987).
4. Chodosh, L.A. *et al.* *Cell* **53**, 11–24 (1988).
5. Rosenfeld, P.J. & Kelly, T.J. *J. biol. Chem.* **261**, 1398–1408 (1986).
6. Chodosh, L.A. *et al.* *Cell* **53**, 25–35 (1988).
7. Short, N.J. *Nature* **331**, 393–394 (1988).
8. Struhl, K. *Nature* **332**, 649–650 (1988).
9. Hope, I.A. *et al.* *Nature* **333**, 635–640 (1988).
10. Sigler, P.B. *Nature* **333**, 210–212 (1988).
11. Shore, D. & Nasmyth, K. *Cell* **51**, 721–732 (1987).
12. Buchman, A.R. *et al.* *Molec. cell. Biol.* **8**, 210–225 (1988).
13. Bram, R.J. & Kornberg, R. *Molec. cell. Biol.* **7**, 403 (1987).

to result in localized spins, antiferromagnetic exchange and semiconducting behaviour.

The $[ABO_2]_x$ blocks in the $n=1-3$ materials modify adjacent $[CuO_2]_x$ layers because the copper ions are no longer surrounded by just four co-planar oxygen ions, but by a square pyramid of five oxygen ions — four in the $[CuO_2]_x$ layer and one from the $[ABO_2]_x$ block. Although this is unlikely to be the cause of the insulating nature, inhibition of superconductivity is possible, because all known copper oxide superconductors contain layers of 'Jahn-Teller' distorted copper ions bonded to either five oxygens (square pyramidal with an elongated vertical bond; for example, $YBa_2Cu_3O_7$) or six oxygens (octahedral with four short and two long bonds; for example $(La,Sr)_2CuO_4$). Because the evidence is limited, this feature cannot with certainty be regarded as essential; however, whereas strontium or barium doping of La_2CuO_4 yields superconductors, similar doping seems ineffective in Eu_2CuO_4 , whose $[CuO_2]_x$ layers and copper coordination

are very much like those of $Ca_{0.86}Sr_{0.14}CuO_2$.

Despite being electrically insulating, the structure and chemistry of $Ca_{0.86}Sr_{0.14}CuO_2$ could give useful indications of the requirements for copper oxide phases to be superconducting, and hence assist the formulation of new materials. Initially, however, the influence of copper valence requires study, and doping experiments aimed at introducing some Cu^{3+} are eagerly awaited. If these materials exhibit superconductivity, their simple structure, involving only two-dimensional $[CuO_2]_x$ layers, will be invaluable for testing theoretical models. □

1. Greaves, C. & Forgan, E.M. *Nature* **332**, 305 (1988).
2. Torardi, C.C. *et al.* *Science* **240**, 631-633 (1988).
3. Sunshine, S.A. *et al.* *Phys. Rev. Lett.* (in the press).
4. Wheatley, J.M. *et al.* *Nature* **333**, 121 (1988).
5. Siegrist, T., Zahurak, S.M., Murphy, D.W. & Roth, R.S. *Nature* **334**, 231-232 (1988).
6. Tarascon, J.M. *et al.* *Phys. Rev. B* (in the press).
7. Cheetham, A.K. *et al.* *Nature* **333**, 21 (1988).
8. Hybertsen, M.S. & Mattheiss, L.F. *Phys. Rev. Lett.* **60**, 1661-1664 (1988).
9. Krakauer, H. & Pickett, W.E. *Phys. Rev. Lett.* **60**, 1665-1667 (1988).

Colin Greaves is in the Department of Chemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Nuclear physics

Thick-skinned lithium isotope

P.G. Hansen

A THICK neutron skin, or 'halo', around neutron-rich nuclei seems to have been identified by a Japanese group, Kobayashi *et al.*¹, working at Lawrence Berkeley Laboratory. The halo, seen on the isotope ${}^9\text{Li}$, could extend to several times the radius of the nuclear core. Because, as one astrophysicist remarked, neutron-saturated nuclei are the closest one can get to having a neutron star in the laboratory, this unexpected finding is likely to influence predictions of the structure and properties of very neutron-rich objects. Such nuclei will also give interesting insights into the structure of other loosely bound quantum systems.

The limiting number of neutrons that can be bound by a given element, the so-called drip line, can be reached experimentally only for the lightest elements. The study of drip-line nuclei has progressed remarkably, largely because of the development of recoil spectrometers that can identify fast fragments from heavy-ion reactions at intermediate energies (see my News and Views article last year²). The experimental programme initiated by Tanihata *et al.*³ takes this technique a step further by observing nuclear reactions caused by the radioactive fragments.

In the experiments (see figure)

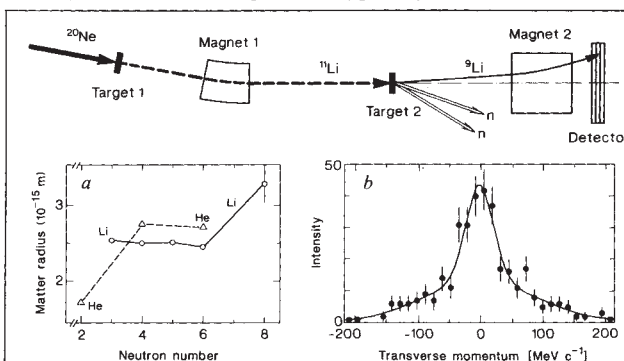
the fast primary beam, here ${}^{20}\text{Ne}$ at 790 MeV per nucleon, reacts in a first target, producing a secondary beam of a single radioactive fragment (here ${}^9\text{Li}$, with only three protons) selected from the debris by magnetic-analysis, charge-detection and time-of-flight techniques. This beam subsequently interacts with a second target and the emerging fragment, ${}^9\text{Li}$, is identified in a spectrometer. The observed transmission of the fragments through the second target (b in the figure) provides an estimate of the nuclear matter radii. Nuclear charge radii typically differ only

slightly from matter radii, but the new results are very striking. Whereas all other nuclei with 3-8 neutrons have radii close to 2.5 fm (2.5×10^{-15} m), the neutron-rich ${}^6\text{He}$, ${}^8\text{He}$ and ${}^9\text{Li}$ have larger radii. Tanihata *et al.*³ interpreted this as evidence of either a large deformation (cigar-shape) or a long tail in the radial matter distribution.

Evidence supporting the second hypothesis comes from a sophisticated experiment on ${}^9\text{Li}$ carried out at the CERN ISOLDE installation by Arnold *et al.*⁴. In this, lithium atoms were polarized in flight by a laser beam and the polarization was detected via the angular asymmetry of the β -radiation (${}^9\text{Li} \rightarrow {}^9\text{Be} + e^- + \nu$). Both the nuclear spin and the magnetic moment were in excellent agreement with the value assigned for the odd proton in the spherical shell model, which suggests that the large radius is attributable to the neutrons alone.

A more direct observation of the halo has now been achieved in the Japanese experiment¹ in which the distribution of the transverse momentum of the ${}^9\text{Li}$ fragment was measured (c in the figure). This distribution has two components: a broad one of full width at half maximum (FWHM) of about 220 MeV c^{-1} (c is the speed of light), typical of that observed in neighbouring nuclei; and a narrow one with a FWHM of 50 MeV c^{-1} . To understand this, note first that the momentum emerging in a fragmentation reaction is essentially that existing in the nucleus before the collision, and second that the spatial and momentum distributions are linked through Heisenberg's uncertainty principle, being Fourier transforms of each other. Hence the narrow momentum component is evidence of a wide spatial component: the neutron halo.

There are several discussions⁵⁻⁸ of the phenomenology of the neutron halo. It is not surprising that the shell model⁵ has difficulties in reproducing the neutron halo. In fact, there is no strong central potential, so essential to the shell model, in ${}^9\text{Li}$, which will not bind a single extra neutron. It can bind two, but with a total binding energy B of only 190 ± 110 keV. This indicates that the neutron-neutron force in the relative S-state (the two neutrons have the same angular momentum) is essential to the stability of ${}^9\text{Li}$ and other drip-line nuclei and revives⁶ Migdal's idea⁹ that the di-neutron, being very close to the bound state, can be stabilized by the attractive interaction of a distant nucleus. Semi-quantitative estimates based on this picture agree with the observed radii and transverse momentum spread, proportional to $\sqrt{B}$ so that the predicted FWHM is 30 ± 10 MeV c^{-1} .



In the experiments of Tanihata *et al.*³ and Kobayashi *et al.*¹, ${}^{20}\text{Ne}$ ion collisions with target 1 produce a shower of ions from which ${}^9\text{Li}$ ions are selected magnetically. The interaction of these neutron-rich fragments with the second target provides two clues to the halo structure: a, an increase in the matter radii² and b, a narrow component in the transverse momentum distribution¹ of the tertiary beam (${}^9\text{Li}$), also selected magnetically.

The large distance of the neutrons from the charged nuclear core suggests^{6,7} that there could be a large intrinsic electric dipole moment because the centre of mass and centre of charge do not coincide. This leads to large cross-sections for Coulomb dissociation in nuclear collisions. These cross-sections are roughly proportional to $1/B$ and reach many barns ($1 \text{ barn} = 10^{-28} \text{ m}^2$) for heavy targets. Estimates agree qualitatively with the measured values⁸ for a lead target. It is interesting that the electric-dipole mode involved is a very soft one. Sum-rule techniques can be used to estimate that the average energy deposited in the dissociation of ^{11}Li into ^9Li and a di-neutron is only around $6B$, or 1 MeV .

Lombard and Maillet propose⁸ that another way to detect the halo structure is to measure the ratio of production of π^- and π^+ in collisions of neutron-rich fragments. This scheme, which seems experimentally feasible, exploits the fact that the surface nucleons are more active in the pion production process than those in the core, so that the halo should enhance π^- production arising from the excitation of neutrons.

The most interesting structural aspect of the halo nuclei is that the neutron-neutron interaction is as strong as or stronger than the interaction of a single neutron with the core. The study of the ensuing neutron-neutron correlation in the wavefunction presents an interesting challenge. One way of tracing this correlation experimentally is to observe in coincidence the neutrons emerging from dissociation, so that one obtains a map of the wavefunction in momentum space.

Structures analogous to the halo nuclei are of interest in atomic and molecular physics. The two valence electrons in negative alkali ions show strong correlations similar to those between the halo neutrons. An even closer analogy is the attraction between neutral rare-gas atoms through van der Waals force which, like the nuclear force, is of limited range. It is conjectured¹⁰ that the attraction between two helium atoms, too weak to make He_2 bound, is strong enough to make He_3 bound at low temperatures. This molecule would be similar to the nucleus ^{11}Li , for which the three-body system ($^9\text{Li}+n+n$) is marginally bound whereas none of the three two-body sub-systems possesses bound states. □

1. Kobayashi, T. *et al.* *Phys. Rev. Lett.* **60**, 2599–2602 (1988).
2. Hansen, P. G. *Nature* **328**, 476–477 (1987).
3. Tanihata, I. *et al.* *Phys. Rev. Lett.* **55**, 2676–2679 (1985).
4. Arnold, E. *et al.* *Phys. Lett.* **197B**, 311–314 (1987).
5. Poppelier, N. A. F. M. *et al.* in *Nuclei Far From Stability: Am. Inst. Phys. Conf. Proc.* **164**, 334–343 (AIP, New York, 1988).
6. Hansen, P. G. & Jonson, B. *Europhys. Lett.* **4**, 409 (1987).
7. Bertulani, C. A. & Baur, G. *Nucl. Phys. A* **480**, 615 (1988).
8. Lombard, R. J. & Maillet, J. P. *Europhys. Lett.* **6**, 323 (1988).
9. Migdal, A. B. *Sov. J. nucl. Phys.* **16**, 238–241 (1973).
10. Macek, J. *Z. Phys.* **D3**, 31–37 (1986).

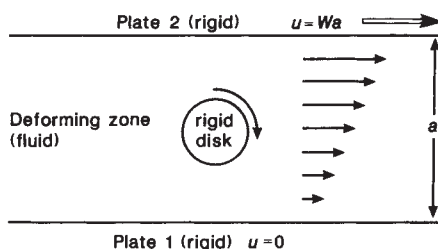
P. G. Hansen is at the Institute of Physics, University of Aarhus, DK-8000 Aarhus C, Denmark.

Continental tectonics

Rock around the clock

James Jackson

GEOLOGICAL strata that have been tilted about a horizontal axis are familiar features of deformed continental regions. It is now clear that large areas of continental crust also rotate about a vertical axis during deformation, although this is harder to demonstrate as there is rarely visible evidence of such rotation in the field. Instead, evidence is found in measurements of the Earth's magnetic field frozen into rocks as they formed. Rotations are revealed by comparing



Plan view of a deforming fluid separating two rigid plates with relative motion $u = Wa$ parallel to the zone. The vertical vorticity of the fluid is W and a rigid disk floating in the fluid rotates with an angular velocity of $W/2$. Palaeomagnetic measurements show that blocks tens of kilometres across can rotate in this way.

palaeomagnetic declinations (the direction of magnetic north at the time of rock formation) within the deforming regions with those found in rocks in the apparently stable regions outside the deforming zone. Theorists and geologists from diverse fields recently discussed the data and models concerning these rotations at the first interdisciplinary meeting* addressing the issue. One of their principal conclusions is that such rotations can often be described by treating the continents as continuously deforming fluids: a statement that will seem strange to geologists and seismologists who know that the continents deform discontinuously by faulting.

Large rotations about a vertical axis were first identified in western North America in the late 1970s (M. Beck, Western Washington University; B. Luyendyk, University of California, Santa Barbara). There had been no predictions that such rotations would be found, nor that they would be so widespread, for example, covering most of the Transverse Ranges; regular, for nearly all those in south-west California are in a clockwise sense, or rapid — up to 90 degrees in the past 10 million years. In particular, thorough geological surveys in southern

California — oil exploration has provided abundant high-quality surface and sub-surface observations — gave no indication of these rotations. Similar rotations have since been found in other areas of young continental tectonic activity, including New Zealand (R. Walcott, Victoria University of Wellington), Israel (Z. Garfunkel, Hebrew University, Jerusalem), Greece, western Turkey and the Andes (C. Kissel and C. Laj, Gif-sur-Yvette), Alaska (R. Coe, University of California, Santa Cruz), and Washington and Oregon (R. Wells, US Geological Survey). At all these sites, the continents are undergoing a component of distributed shear parallel to the length of the deforming belts.

At the same time, and independently, new models of continental deformation were being developed. It was known that the continents do not deform in the same way as the oceans: continental seismicity (and hence faulting) is far more widespread and is not confined to narrow, linear belts. The familiar rules of plate tectonics, which are so useful in studies of the oceans, seem not to apply to the continents because the kinematics within the wide deforming zones are not easily related to the motions of the stable regions that bound them. Within the deforming zones it is more useful to abandon the concept of rigidity, central to plate tectonics, and describe the observed motions as those of a deforming fluid in which there are no strain discontinuities at all.

Such models successfully describe the long-wavelength strain in continental terrains (Walcott; P. England, University of Oxford). But they also predict, within the errors of measurement and dating, the observed sense and rate of rotations about a vertical axis in terrains of distributed deformation. This depends on the assumption that such rotations represent the behaviour of rigid blocks responding to the vorticity of the fluid (D. McKenzie and myself, University of Cambridge; see figure). Although theory predicts that the shape of the blocks affects their rotation (S. Lamb, Victoria University of Wellington), the observed rotations in New Zealand, Peru, California, Nevada and Greece are those expected for a circular disk, suggesting that the blocks are roughly this shape.

In retrospect, the success of continuum models of continental deformation at long wavelengths is less surprising: it is known that only the top 10–20 km of the lithosphere deforms in a brittle discontinuous fashion (in earthquakes) and that the lower 100 km presumably deforms by

* Palaeomagnetic Rotations and Continental Deformation
NATO Advanced Research Workshop, Loutra Edipsos, Greece, 8–12 May 1988.

distributed creep or flow. Thus at a length scale large compared with the thickness of the brittle upper crust it is reasonable that deformation is dominated by the distributed flow in the lower crust and mantle lithosphere.

But a paradox remains: the deformation may appear continuous at large length scales, but we know that in reality it is discontinuous in the upper crust and occurs on faults. How do the faults move to accommodate the flow below, and what constraints does this requirement put on their possible geometry and kinematic evolution? In particular, how are the faults able to accommodate rotations about a vertical axis? Attempts to relate rotations to faulting in Greece (McKenzie and myself), New Zealand (Lamb), Israel (Garfunkel; A. Nur, Stanford University), California (Luyendyk; C. Nicholson, University of California, Santa Barbara) and Tibet (P. Molnar, Massachusetts Institute of Technology) do not address the question of what happens at

depth that allows the transition from distributed flow to discontinuous faulting. Theoretical arguments and analogue experiments (J.P. Brun, Rennes University) suggest that in such regions the fault geometries change rapidly with time and that a knowledge of slip vectors on the faults is required to see how they accommodate the deformation: both of which will make the analysis of older terrains difficult.

Rotations about a vertical axis have obvious practical implications for those attempting palaeogeographic or palaeocurrent analyses: recording a strike direction is often the first field technique geologists are taught, but is starting to look like one of the most difficult geological observations to interpret. Not for the first time, palaeomagnetism has raised questions that are likely to occupy tectonic geologists for some time. □

James Jackson is in the Bullard Laboratories, University of Cambridge, Cambridge CB3 0EZ, UK.

Plant molecular biology

Plant development and ribozymes for pathogens

Virginia Walbot and George Bruening

DRAMATIC developments in the use of catalytic sequences from satellite RNAs to cleave mRNA specifically, and so to control gene expression, provided one of the highlights of three related meetings*. Another topic that aroused great interest, which we will describe first, was the beginning of a fundamental insight into the mechanism of hormone action in plant development.

Stress hormone in development

The hormone abscisic acid (ABA) has a critical role in plants not only in the response to water stress, but also in processes central to the normal life cycle. The most important of these is the desiccation of the embryo and surrounding maternal tissues to produce a metabolically quiescent seed. But almost nothing is known about the mechanisms whereby plants respond to this important hormone. So work at the University of Florida and at CSIRO, Canberra indicating that this question may yield to a molecular biological approach aroused considerable interest at Steamboat Springs, where the response to ABA was reported to be associated with the induction of specific proteins both in water-stressed plants and in developing embryos.

The developmental programme of a plant embryo consists of three phases: the elaboration of the embryonic organs; the accumulation of storage reserves; and desiccation of the embryo and surrounding maternal tissues to produce the seed. During the turning point between stages two and three, the embryo must acquire the ability to withstand a reduction in water content from about 85 per cent to 10 per cent of fresh weight. In other plant parts, such severe desiccation kills cells.

Among the experiments over the past decade that have demonstrated a link between ABA and the progression of embryos from stage two to three are studies on viviparous mutants of maize. Viviparous embryos fail to desiccate at stage three and instead imbibe water from surrounding tissues and germinate within the fruit. (This syndrome is common in cultivated grapefruit in which roots emerging from the seed are pickled in the highly acidic flesh; not surprisingly, viviparous mutations are generally lethal.)

In most mutants vivipary seems to result from a lack of ABA: thus viviparous mutants of maize can be coaxed in culture to develop quasi-normally by the addition of exogenous ABA (Robichaud, C. S. *et al.*, *Dev. Genet.* **1**, 325–330; 1980). The relatively rare mutants in which vivipary occurs despite normal levels of abscisic acid in the tissues may therefore reflect mutations affecting the response to the

hormone. The recent investigations of D. McCarty (University of Florida) on one such mutant, *vivipary 1* (*Vp1*), now open this hypothesis to test. He has isolated the *Vp1* allele by transposon tagging and has preliminary evidence for a complex pattern of transcription in embryonic organs. Mutant alleles lack one or more of the transcript size classes. The *Vp1* phenotype suggests that the wild-type allele encodes either an ABA receptor in embryo tissue, or some other part of the signal-transduction chain linking hormone stimulation to alterations in gene expression. In either case, elucidation of *Vp1* function should prove exciting.

An important link connecting the presence of ABA to embryo survival during desiccation was also reported. Using a collection of complementary (c)DNAs encoding ABA-induced proteins in barley aleurone as a starting point, P. Chandler (CSIRO, Canberra) has discovered a new class of proteins in grasses that accumulate in all organs subjected to water stress. Furthermore, accumulation of these proteins seems to protect plant tissues from death during desiccation. The best-studied example of this new class of proteins, from barley, has a relative molecular mass of 22,000 (22K). It consists of a repetitive amino-acid sequence, is extremely hydrophilic and has the unusual property of not precipitating after boiling. Corn contains a very closely related 25K protein — one region of 50 amino acids matches perfectly with a region of the 22K barley-protein sequence. Also of interest is the report by Gomez *et al.* on page 262 of this issue that a small protein of 15K with an unusually high glycine content is induced in maize by desiccation or ABA treatment.

The 22K protein is induced in barley embryos by 25–50 nM ABA, which is only a 3–4-fold increase over the constitutive hormone level. Moreover, the mRNA encoding the protein can be induced 100-fold by mild desiccation. When water stress is relieved, the mRNA for the 22K protein is rapidly degraded and its synthesis ceases. Treatment of well-watered tissues with ABA also induces the 22K protein. Furthermore, seedlings pretreated with ABA for 24 h to induce substantial amounts of the 22K protein can survive rapid desiccation (50 per cent decrease in fresh weight in four hours) whereas uninduced tissues die.

During normal development it is hypothesized that the 22K protein is induced by ABA and prepares embryos for either the desiccation or cellular disruption that might occur during rehydration. Evidence supporting this hypothesis comes from additional studies examining the expression of the related 25K protein in maize. Chandler and M. Crouch (Indiana University) excised viviparous embryos unable to synthesize ABA and placed them in culture with

*The UCLA joint meetings on the Molecular Basis of Plant Development and Plant Pathology, Steamboat Springs, Colorado, 26 March–1 April, 1988; the UCLA meeting on the Molecular Biology of RNA, Keystone, 2–9 April, 1988.

sufficient water or under water stress. During desiccation, the viviparous embryos fail to accumulate the mRNA encoding the 25K protein, whereas normal sibling embryos synthesize large amounts of this mRNA. Thus, the induction of the desiccation protein depends on the ability to synthesize or respond to ABA. Future studies with water stress in leaves, including plant mutants impaired in ABA metabolism, will be required to determine if the desiccation protein is unique to embryos or young seedlings, or whether it is induced during water stress at any developmental stage.

Designer ribozymes

Two years ago G. A. Prody *et al.* (*Science* **231**, 1577–1580; 1986) reported that dimeric and trimeric forms of satellite RNA of tobacco-ringspot virus spontaneously and specifically cleave to release the 360-base monomeric RNA. Since then about a dozen RNAs, including both polarities of several plant-virus satellite RNAs, one viroid, an RNA from newt and the hepatitis delta virus, have been shown to undergo autolytic cleavage. Most of the autolytically processing sequences fit a consensus secondary structure known as a 'hammerhead' or 'T' structure. Figure 1 illustrates the proposed secondary structure of the satellite RNA of tobacco ringspot virus immediately after autolytic processing. The new ends created in the reaction are a 5'OH and a 2',3'-cyclic phosphodiester, showing that autocatalytic processing is a phosphotransfer reaction.

There are few restraints on the sizes or sequences of the stems in the self-cleaving structures (Forster, A. C. & Symons, R. H. *Cell* **49**, 211–220; 1987). The consensus structure in Fig. 2, modified from a report by Forster and Symons (*Cell* **50**, 9–16; 1987), shows how few invariant sequences are found. Base-paired (N) and not base-paired (X) nucleotides are indicated. Although loops are found in intramolecular cleavage structures (Fig. 1), they are not a requirement for the reaction: O. C. Uhlenbeck (*Nature* **328**, 596–600;

1987), demonstrated that a cleavage complex lacking loops can be assembled from two short (19-mer and 24-mer) partially complementary oligoribonucleotides, the synthetic 19-mer catalytically cleaving the 24-mer at a specific phosphodiester bond (see also Koizumi, M. *et al.* *FEBS Lett.* **228**, 228–230; 1988).

One challenge to the universality of the hammerhead model, however, has been that, in contrast to the viral satellite RNAs, the sequences of self-cleaving viroid and newt RNAs look unlikely to allow formation of a stable stem III when folded into a hammerhead. Symons's group reported evidence, which appears on page 265 of this issue, that salvages the hammerhead model. When two self-cleaving domains are folded together a stable structure can be formed essentially because the weak stem-III-forming sequences cooperate. This scheme is borne out by an experiment showing that

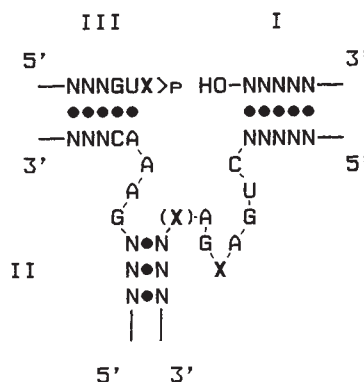


Fig. 2 Consensus sequence for an autolytically processing sequence, modified from Forster, A. C. & Symons, R. H., *Cell* **49**, 211; 1987.

the self-cleavage reaction of an RNA predicted not to form a stable single hammerhead is in fact bimolecular. The same scheme could also apply to covalently linked dimers, which would be intramolecular and thus more efficient.

The inefficiency of a bimolecular reaction could be beneficial to the viroid: it would help prevent unnecessary cleavage of circular monomers presumed to be the templates for the 'rolling circle' replication of plus- and negative-strand RNAs of viroids, but still allowing efficient cleavage of the RNA products of replication when they have more than one self-cleaving domain and can thus fold into double hammerheads by intramolecular interactions. In experiments in which sequences critical for RNA cleavage are placed on two different molecules, the reaction complex may involve the association of four RNA molecules.

W. Gerlach and J. Haseloff (CSIRO, Canberra; and *Nature*, in the press) have exploited this understanding of the structural relationships of ribozyme and substrate to design ribozymes to cleave specific targets. Crucial to their success is

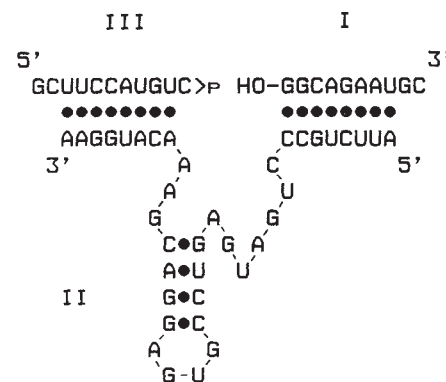


Fig. 3 Endonucleolytic sequence acting on chloramphenicol acetyltransferase mRNA.

the relative insensitivity of the autolytic processing reaction to insertion in the loop of stem I. Figure 3 shows their clever application of this observation in the development of endoribonucleolytic oligoribonucleotides directed to mRNA-size molecules. Three GUC-containing sequences of the mRNA-chloramphenicol acetyltransferase (CAT) are targeted for cleavage by three different oligoribonucleotide ribozymes. Note that GUX is part of the consensus of stem III and that cleavage occurs next to X (Fig. 2). Each ribozyme has two 8-nucleotide regions of complementarity to the CAT mRNA sequence, involving sequences sufficient to form stems I and III, plus the satellite tobacco-ringspot virus RNA sequences in region II. *In vitro*, each oligoribonucleotide cleaves the mRNA at the expected site. Tests of efficiency *in vivo* are now in progress.

Because the substrate requirements are so minimal (GUX in the mRNA), ribozymes capable of cleaving cellular mRNAs should be readily designed simply by including sufficient complementarity in the ribozyme to ensure specificity. Ribozymes targeted to plant viral RNAs, in which elimination of the replicating form of the virus suppresses infection, may be among the first applications of this new technology. Destruction of mRNAs from DNA viruses should also be possible.

Ribozymes should provide an alternative to antisense constructs for probing the role of genes, particularly in organisms lacking strong genetics. The phenotypes associated with the lack of a functional product can be determined if the product can be suppressed. Eliminating a required mRNA, either constitutively or at specific times by using an inducible ribozyme, promises to be more efficient than the antisense method because the RNA cleavage reaction is catalytic.

Virginia Walbot is in the Department of Biological Sciences, Stanford University, Stanford, California 94305 and George Bruening is in the Department of Plant Pathology, University of California at Davis, California 95616, USA.

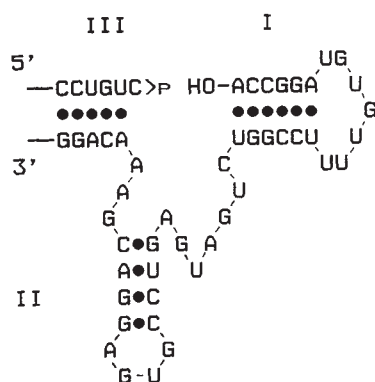


Fig. 1 Postulated autolytic processing sequence of satellite tobacco-ringspot virus RNA.

Greenhouse effect

Methane linked to warming

Ralph J. Cicerone

METHANE is an important chemical in the atmosphere; it controls numerous chemical processes and species in the troposphere and stratosphere and its infrared spectrum makes it a strong greenhouse gas. There is compelling evidence that the concentration of methane has increased globally at a rate of about 1 per cent per year since 1978¹ (and probably since the early 1950s²) to about 1.70 parts per million (p.p.m.) in 1988. This rapid contemporary increase began 150–200 years ago, when the concentration was 0.65–0.70 p.p.m. according to analysis of air trapped in dated polar ice cores^{3–5}. Two new experimental studies by Stauffer *et al.*⁶ and Raynaud *et al.*⁷ extend the ice-core record back to 100,000 years before present (BP)⁶ and 161,000 BP⁷, through the last two major glacial periods, 18,000 BP and 150,000 BP. During glacial maxima, the methane concentration fell to 0.35 p.p.m. whereas during interglacial times it rose to 0.65 p.p.m., very near modern pre-industrial values. Thus, the present concentration and that projected for the future are well above any in charted Earth history.

Stauffer *et al.*⁶ analysed 24 ice-core samples from Antarctica and Greenland to find the methane pattern during the last glaciation and through the transition to the warmer interglacial. Their Antarctic data extend back to 51,000 BP, the Greenland data back to 100,000 BP. Generally, during the last glaciation, about 20,000 BP, the methane concentration was about 0.35 p.p.m., increasing to about 0.65 p.p.m. by 14,000 BP. Preceding the last glacial period, it was 0.45–0.50 p.p.m.

A similar pattern emerges for the preceding glaciation through analysis of

the remarkable Vostok ice core from Antarctica. Raynaud *et al.*⁷ extracted methane from 27 core sections from the 155,000-BP glacial period, the following interglacial (about 130,000 BP) and the intervening transition period. During the glaciation the average methane concentration was 0.34 p.p.m., rising to 0.46 p.p.m. during the transition and 0.62 p.p.m. during interglacial times. Clearly, the natural methane background concentration during warm times is 0.6–0.7 p.p.m. but is a factor of two lower during glacial epochs. There are even indications from Greenland cores that the concentration was low during a cold subinterval (Younger–Dryas period) of the present interglacial, and relatively high during a warm period of the last glaciation⁸. One measurement conflicts with this picture; it is the only previous methane concentration published⁹ for a glacial period (27,000 BP), measured as 0.65 p.p.m..

The analytical skill needed for these measurements deserves recognition. Even after a core is obtained and dated successfully, subsequent contamination, the gas-extraction method and chemical, biological or physical factors could all alter the methane content of the trapped gas^{3–7}. All the indications are that the methane found in the ice cores does truly represent the composition of the atmosphere on the date of firn closure, when the compacting ice enclosed the gas and became impermeable.

Now that we know that the concentration of methane doubled between the lows and interglacial highs, and halved again, we can ask why and use this information to elucidate several components of the

Earth's climate and biogeochemistry. The rate of microbial methanogenesis generally increases with temperature. Also, because the dominant methane sources are on land and not in the oceans¹⁰, these sources should be reduced when wetland soils are ice-covered or frozen, sealed off from the atmosphere. Thus our general knowledge of natural sources of methane is consistent with the new ice-core data.

But the principal sink of atmospheric methane, reaction with hydroxyl (OH) radicals in the troposphere, should also increase with temperature for two reasons. First, the atmospheric OH concentration generally increases with that of H₂O, rising with temperature. Second, the rate constant for the reaction of CH₄ with OH also increases with temperature. But the OH concentration also depends strongly on the amounts of ozone and nitrogen oxide present and the intensity of ultraviolet light. Without a knowledge of these, it is not possible to predict how atmospheric methane destruction rates respond to temperature perturbations.

A warmer Earth with more wetlands would provide more anoxic sites for anaerobic methanogenic bacteria. Overall rates of biological carbon cycling could also be larger. Ice-core CO₂ data have shown a pattern similar to that of methane⁹. Both the oxic and anoxic paths of carbon cycling could be slowed during glaciation. The extent to which the decrease (or increase) of CO₂ concentration actually causes glaciation (or deglaciation) is still debated, but we know now that the greenhouse effects of altered CH₄ must accompany and add to those of CO₂ changes. It is dangerous to argue, though, that other indirect effects would ensue. For example, increased CH₄ need not have led to more tropospheric ozone (another greenhouse gas), as has been proposed⁷, because in past epochs there may not have been sufficient atmospheric NO_x to allow this¹⁰. But it is sobering to learn that the methane concentration, like those of carbon dioxide and chlorofluorocarbons, has already increased to values above those of at least the past 160,000 years and that human activities are clearly involved in these global changes^{4,8}. □

1. Blake, D.R. & Rowland, F.S. *Science* **239**, 1129–1133 (1988).
2. Rinsland, C.P., Levine, J.S. & Miles, T. *Nature* **318**, 245–249 (1985).
3. Craig, H. B. & Chou, C. C. *Geophys. Res. Lett.* **9**, 1221–1224 (1982).
4. Rasmussen, R. A. & Khalil, M. A. K. *J. geophys. Res.* **89**, 11599–11605 (1984).
5. Stauffer, B., Fischer, G., Neftel, A. & Oeschger, H. *Science* **229**, 1386–1388 (1985).
6. Stauffer, B., Lochbrunner, E., Oeschger, H. & Schwander, J. *Nature* **332**, 812–814 (1988).
7. Raynaud, D., Chappellaz, J., Barnola, J., Korotkevich, J. S. & Lorius, C. *Nature* **333**, 655–657 (1988).
8. Ehhalt, D. H. *Tellus* **26**, 58–70 (1974).
9. Lorius, C. *et al.* *Nature* **316**, 591–596 (1985).
10. Volz, A. & Kley, D. *Nature* **332**, 240–242 (1988).

Ralph J. Cicerone is at the National Center of Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA.



100 years ago

ON Thursday, the 12th inst., the anniversary meeting of the Sanitary Institution of Great Britain was held in the theatre of the Royal Institution. The Chairman, Mr. Edwin Chadwick, in opening the proceedings, claimed credit for the Sanitary Institution of Great Britain and like institutions for a large proportion of the reduced death-rate of the metropolis, which was now 14 in 1000. London in that respect compared very favourably with other places, the death-rate in Paris being 27, Vienna 30, and St Petersburg 40. Dr. B. W. Richardson delivered an address on "The Storage of Life as a Sanitary Study." He began by referring to

instances of long life in lower animals and in man. These, he said, by some peculiar process as yet but little investigated, held life as a long profession, and to this faculty he applied the term "The Storage of Life." The problem which the lecturer placed before the society was stated as follows:—Certain proofs of the power of the human body to lay or store up life to a prolonged period are admitted. What are the conditions which favour such storage, and how can we promote the conditions which lead to it? He stated the conditions in the following order, hereditary qualifications, the virtue of continence, maintenance of balance of bodily functions, perfect temperance, and purity from implanted or acquired diseases. In dealing with all-round temperance, he showed that whatever quickened the action of the heart beyond its natural speed and force was a stimulant, and in proportion to the unnatural tax inflicted by stimulation there was a reduction in the storage of life.

From *Nature* **38**, 276; 19 July 1888.

Immunology

Cell-cell interactions in the antibody response

Anthony L. DeFranco

MANY immunological observations made over the past two decades have led to the formulation of a scheme in which the T cells that 'help' B cells make antibody are themselves activated by antigen bound to the surface of the B cell. This reciprocal activation in its simplest form is however challenged by O. Lassila and colleagues, who report on page 253 of this issue¹ that it cannot take place unless the T cells have first been activated in a separate process that does not involve B cells.

Activation of resting B cells starts with binding of antigen to cell-surface forms of immunoglobulins. For most protein antigens, the antibody response of B cells also depends on signals received from helper-T cells. These do not recognize antigen in its native form but instead recognize a complex of an antigen-derived oligopeptide and a major histocompatibility complex (MHC)-encoded molecule on the surface of another cell, the antigen-presenting cell. The pioneering experiments of Mitchison² showed that for helper-T cells and B cells to collaborate, they must recognize parts of the same molecule.

This and subsequent experiments argue compellingly that the helper-T cell needs to be physically close to the B cell. Helper-T cells secrete B-cell growth and differentiation factors, such as interleukins 4, 5 and 6, and the need for close contact may reflect a requirement for the higher concentrations of these factors present at their release site. There is now also good evidence that helper-T cells can direct secretion of interleukins towards antigen-presenting B cells^{3,4}, which would increase this effect. Recent experiments^{5,6} show that when a B cell presents the appropriate complex of antigen fragment and MHC protein to a helper-T cell, a stable interaction results, inducing the T cell to produce interleukins⁷⁻⁹ that activate the B cell.

How does the B cell acquire the antigen for presentation? If a B cell captures an antigen through surface immunoglobulin rather than by pinocytosis, the antigen is presented more than a thousand times as efficiently^{10,11}. Thus the observation that secondary antibody responses require B cells and helper-T cells to recognize the same antigen¹ presumably reflects the following series of events: preferential uptake of one particular antigen by the antigen-specific B cell, internal processing of the antigen to peptides, binding of the peptides to MHC molecules of that B cell, and interaction of that B cell with a helper-T cell specific for the same antigen. The

T cell then delivers the regulatory signals that the B cell requires to make a vigorous antibody response (see Fig. 1).

The issue of whether B cells can present antigen to resting, unprimed T cells and thereby initiate the whole process has been difficult to address. Most of the experiments have used B cells to present antigen to T-cell hybridomas, to cloned T-cell lines, or to *in vivo* primed T cells, but not to resting, unprimed T cells (reviewed

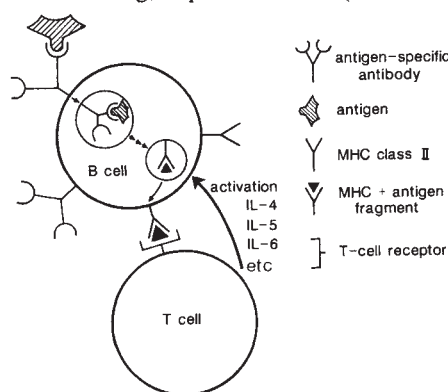


Fig. 1 Current view of events involved in B cell-T cell interaction. First, the antigen in its native form binds to membrane immunoglobulins on the surface of the B cell. Then the antigen is internalized by endocytosis and degraded to oligopeptides, which bind to MHC class II proteins, probably within the cell. Next the peptide-MHC class II protein complex goes to the cell surface, where it is recognized by an antigen-specific helper-T cell. The interaction between the T-cell antigen receptor and the peptide-MHC class II protein complex leads to a stable interaction between the T cell and the B cell, and to directed secretion by the T cell of interleukins that are activation signals for the B cell.

in ref. 9). Aside from the question of the capability of B cells to prime resting T cells, there is also the issue of probability. Because the frequency of antigen-specific lymphocytes is low before the primary immunization, it may be difficult for two rare cells to find each other. The immune response includes a phase of vigorous clonal expansion of antigen-specific B cells and T cells, so this problem diminishes later in the response. Thus, it is not clear under what circumstances B cells present antigen to helper T cells *in vivo*, or how this relates to the requirement for direct interaction between these cells.

In this issue, Lassila, Vainio and Matzinger¹ report an elegant experimental system that nicely addresses these issues. In chickens, it is possible to ablate B-cell precursors shortly after birth but to leave intact the precursors of T cells,

macrophages and other cells. Introduction of B-cell precursors from another neonatal chicken results in the establishment of a stably chimaeric bird that has T cells and macrophages expressing one set of MHC alleles and B cells expressing another. The different alleles of MHC are sufficiently polymorphic that a T cell that recognizes an antigen fragment bound to one MHC allelic product will not recognize the same fragment bound to the product encoded by a different allele. This is a property of T-cell antigen recognition referred to as MHC restriction. How does this change in MHC molecules between B cells and the rest of the organism affect T-cell-B-cell collaboration? The effect is dramatic. T-cell-dependent antibody responses, both primary and secondary, are completely deficient. The B cells are fully functional, however, because T-cell-independent antibody responses are normal. Thus, the problem must lie in either the T cells themselves or in their ability to interact with B cells expressing different MHC molecules.

It would be reasonable to expect that this defect is due to an absence of helper-T cells able to collaborate with the introduced B cells. Previous experiments indicated that T cells preferentially recognize antigen combined with the MHC molecules they express, rather than with those they do not express. This preference results from a poorly understood positive selection event in the thymus during the maturation of T-cell precursors. The chimaeric birds should express only the endogenous MHC alleles in the thymus and not the alleles expressed by the B cells, and there should, therefore, be very few helper-T cells capable of recognizing antigen plus B-cell-type MHC molecules.

Surprisingly, the number of such T cells is sufficient for normal antibody production. Introduction of irradiated spleen cells (presumably including functional antigen-presenting cells other than B cells, for example, macrophages and dendritic cells) of the same MHC type as the B cells fully restored the antibody response. This indicates that the helper-T cells are present but are unable to participate unless non-B antigen-presenting cells of the donor type are added. The macrophages and dendritic cells of the recipient MHC type are always present and are functionally equivalent to the introduced spleen cells, except for their differences in MHC molecules. It seems these cells must be needed to provide a crucial antigen-presenting function.

The most likely explanation is that a non-B antigen-presenting cell is required to activate the resting helper-T cell initially. Then the helper-T cell finds the antigen-specific B cell and recognizes the same antigen fragment with MHC complex that it recognized earlier on the surface of the introduced non-B cell (Fig. 2).

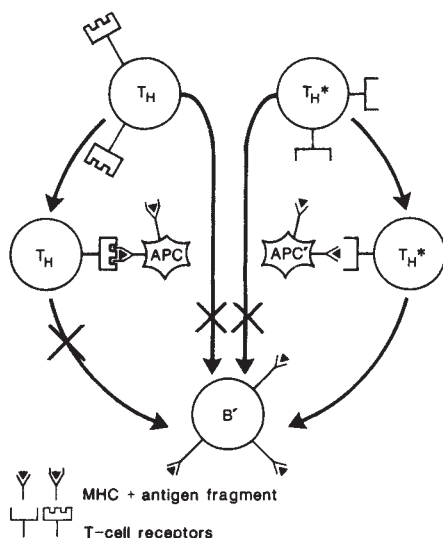


Fig 2. Some resting helper-T cells (left, T_H) recognize an antigen fragment in combination with a self (recipient type) MHC molecule, expressed on the endogenous antigen-presenting cell (APC). These T cells cannot recognize the antigen fragment bound to the non-self (donor type) MHC molecules, expressed by the B cell of the chimaeric chicken (B^*) and thus cannot participate in the T cell-B cell interaction required for antibody production. Other helper-T cells (right, T_H^*) recognize antigen fragment in combination with MHC of the donor type, expressed by B^* and by APC, the introduced non-B antigen-presenting cell. Without APC, no antibody response occurs; when APC is added, a good response ensues. This argues for two-step activation: initial activation of the resting helper-T cell by antigen presented by APC, followed by an interaction of the helper-T cell with an antigen-specific B cell. T_H^* then provides the helper function to B^* , allowing it to become an antibody-secreting cell. Previous work on the ability of T cells to respond to antigen presented by cells expressing differing MHC alleles suggested that the number of T_H^* would be much less than the number of T_H . This does not seem to be true in this system. Paths of T cell-B cell interactions that seem not to occur are marked X.

This recognition establishes the direct T-cell-B-cell interaction needed for antibody production. This strongly argues that, for an antibody response, the helper-T cell must recognize antigen presented by different types of cells, first a non-B cell and then the antigen-specific B cell that is the recipient of the regulatory factors released by the T cell.

It is not clear whether the requirement of a non-B cell for the first antigen-presentation step reflects an actual inability of B cells to activate resting T cells, as the authors conclude in their title, or whether the initial step leads to expansion of the resting T cells, which is needed to increase the probability that they find the appropriate B cells. Interestingly, recent experiments with mice depleted of B cells by neonatal treatment with anti-IgM antibodies have demonstrated that priming of lymph-node T cells requires antigen-specific B cells⁹.

These results seem to be inconsistent with the results of Lassila *et al.*¹. There are, however, several possible reasons why

these two types of experiment could have different outcomes. The first and least satisfying is that chickens and mice differ in the priming requirements of their resting T cells. An alternative is that two types of antigen-presenting cells, a B cell and a non-B cell, may both be required for T-cell priming. Although B cells are needed to prime lymph-node T cells, they may not be sufficient — another type of antigen-presenting cell may be needed as well. This could represent a limitation in the number of the non-B antigen-presenting cells, so that they can initiate but not sustain the response. According to this view, antigen presentation by B cells would be necessary to allow continued replication of antigen-activated T cells. The limitation in the number of the non-B antigen-presenting cells may be greater in the lymph nodes than in the spleen, as B cells are not required for T-cell priming in the spleen, whereas they are in the lymph nodes⁹.

Another system has also recently been developed for studying the nature of the cells that present antigen to T cells to get various immune responses *in vivo*. van Ewijk *et al.*¹² have introduced an I-E class II MHC transgene into the germ line of I-E mice. By deleting various upstream DNA elements that control the transcription of the transgene in a cell-type-specific manner, they have created mice that express I-E molecules on some cell types but not on others. The immunological properties of these mice are surprisingly complicated and further work is required to understand them better.

The dramatic findings reported by Lassila *et al.* may, however, help to explain why some of the transgenic mice were unable to make antibody responses that are restricted to the I-E transgene, for instance, mice that express I-E in B cells but not in macrophages. In any case, the new ability to examine the relative importance of different types of antigen-presenting cells for various *in vivo* T-cell immune responses has already clarified our picture of how T-cell-dependent antibody responses occur, and promises further insights in the near future. □

1. Lassila, O., Vainio, O. & Matzinger, P. *Nature*, **334**, 253–255 (1988).
2. Mitchison, N.A. *Eur. J. Immunol.* **1**, 18–27 (1971).
3. Kupfer, A., Swain, S.L. & Singer, S.J. *J. exp. Med.* **165**, 1565–1580 (1987).
4. Poo, W.J., Conrad, L. & Janeway, C.A., Jr *Nature* **332**, 378–380 (1988).
5. Kupfer, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 6080 (1986).
6. Sanders, V.M. *et al. J. Immunol.* **137**, 2395–2404 (1986).
7. Chesnut, R.W., Colon, S.M. & Grey, H.M. *J. Immunol.* **128**, 1764–1768 (1982).
8. Ashwell, J.D. *et al. J. exp. Med.* **159**, 881–905 (1984).
9. Ashwell, J.D. *J. Immunol.* **140**, 3697–3700 (1988).
10. Rock, K.L. *et al. J. exp. Med.* **160**, 1102–1113 (1984).
11. Lanzavecchia, A. *Nature* **314**, 537–539 (1985).
12. van Ewijk, W. *et al. Cell* **53**, 357–370 (1988).

Anthony L. DeFranco is in the Departments of Microbiology & Immunology and Biochemistry & Biophysics, and the George Williams Hooper Foundation at the University of California, San Francisco, California 94143, USA.

Daedalus

Sounding brass

DAEDALUS once devised a clever but rather clumsy metal detector. A magnetic powder suspension was sprayed into the subject's ears, coating his eardrums. He could then directly hear the a.c. field of a nearby search coil. Metal in the vicinity changed the character of the note, and could be tracked down by seeking the maximum effect. This could be rather inconvenient. A dropped coin, for example, had to be tracked on all fours, with one magnetized ear close to the carpet.

So Daedalus is refining the notion. He is fitting the electronics in a special helmet, from which the search coils jut on outriggers. They are placed so that their fields just cancel at the wearer's ears, which form the null points of the search field. Normally, therefore, the wearer hears nothing. But metal in the vicinity distorts the field, in effect superimposing on it a perturbation centred on the metal. And this, with his magnetically sensitized ears, the wearer hears. He even senses it as coming from the metal. His two ears perceive the spatial distribution of the field-perturbation as a sound, and locate its source by reflex stereoperception. He can track it down like any other sound-source.

This ingenious arrangement is completed by sensitive field detectors at the wearer's ears. They amplify and (for radio-frequency search fields) demodulate the local field for his magnetized eardrums to sense, thus increasing their effective range and sensitivity to metal. His normal hearing, of course, is unaffected.

The wearer of Daedalus's metal-hearing helmet will enter a rich new sensory world. Fallen keys and lost money, wiring in the walls and treasure under the floorboards will chant or chirrup to him at the search frequency, each from its own location. His 'cocktail-party effect' sense of audio-discrimination will let him select any one of them, and track it down. Composite objects like nail clippers or landmines, with metal-to-metal rectifying junctions in their structure, will sing with a range of higher overtones as well. With a swept-radio-frequency source bearing a corresponding audio modulation, each metallic chorister will announce its own identity still more clearly, through its personal set of electromagnetic resonances.

Thus the security guard will soon learn the sonic nuances of hidden revolvers and hand grenades, while the mechanic will seek the more innocent voices of pliers, wrenches and hacksaws misplaced in the clutter of his bench. Both could identify their friends by the timbre of their tooth fillings, and navigate round a darkened house by using door handles, stair rods or plumbing as radio beacons. David Jones

Is HIV the causative factor in AIDS?

SIR—As a participant at the forum in which Peter Duesberg faced his critics (*Nature* 332, 574; 1988), it is my opinion that the Byzantine symptomatology of AIDS is unlikely to be simply accounted for by immunodeficiency supposedly caused by the retrovirus termed HIV.

This is based on experience with the natural spread of another retrovirus, avian leukosis virus (ALV), assumed since 1910 to be responsible for visceral lymphomatosis (VL) in chickens (H. Rubin *et al.* *Virology* 17, 143; 1962). I found that 25% of about 600 chickens in a susceptible flock were congenitally infected with ALV, and almost every cell actively produced virus throughout a bird's lifetime. The other birds became infected by contact after hatching and developed antibodies.

Despite massive and continuous active infection of the congenitally infected chickens, their growth and appearance was the same as that of the other birds. Only 15% of the congenitally infected birds developed VL and only 3% of the contact-infected birds, about the same percentage that develop VL in the absence of ALV infection. Although the virus may play a role in accelerating the onset of the disease, clearly there are other factors involved in its occurrence, including life history and genetic makeup. In an ordinary commercial flock only 3% were congenitally infected, with most of the rest becoming infected by contact. If the occurrence of VL in the infected population of a commercial flock is as low as that in our experimental flock, the overall occurrence of VL would be no higher than in an uninfected flock and the role of the virus, in the absence of other factors, would be moot.

I have found no convincing evidence in the AIDS literature or at the forum that HIV is more cytopathic *in vivo* than ALV. Even were HIV more cytopathic, it is unlikely to account directly for the immune deficiency in view of the tiny fraction, less than 1 in 10,000, of lymphocytes actively infected. The protagonists of HIV causation of AIDS at the forum discounted the significance of other factors prevalent in the high-risk groups, such as multiple infections with other viruses and bacteria, and the debilitating effects on the immune system of repeated receptive anal intercourse, intravenous drug abuse and heavy antibiotic consumption. These practices have greatly intensified in the United States with the advent of gay liberation and the drug epidemic.

The annual AIDS morbidity of 5% among those with HIV antibody in high-risk groups predicts 50,000 to 75,000 new cases per year among the 1–1.5 million US antibody positives, which is 3–5 times

higher than the actual annual occurrence. In fact, it is possible that AIDS morbidity is very low in the absence of 'cofactors', and might approach zero in the healthy heterosexual population, as Duesberg claims. In a complex syndrome like AIDS it is difficult indeed to disentangle cause from correlation without much more data on disease incidence among HIV-positive individuals in low-risk groups. The usual counter argument to any question of the primeval role of HIV in AIDS is the citation of cases induced by transfusion; but this has to be considered alongside the absence of the disease in over 750 accidental needle sticks with AIDS-contaminated blood. In short, the picture is far from simple and it is a disservice to base our health practices solely on that picture.

HARRY RUBIN

Department of Molecular Biology,
University of California, Berkeley,
California 94720, USA

Fossil carbon sources of atmospheric methane

SIR—In support of a recent report¹ that about 32% of atmospheric methane is derived from fossil carbon, we would like to suggest two petroleum-related sources that have not received sufficient attention.

The first is the underwater venting of waste or non-commercial quantities of natural gas from offshore production platforms — estimated to be responsible for up to 70% of the gases being flared or vented offshore in the US². Underwater venting was common in the 1970s in the outer continental-shelf (OCS) region of the US, amounting to a few per cent of the total gas produced. *Sea Technology*³ gives a figure of 60 million MCF (1,000 cubic feet) in the OCS region in 1973. One may thus calculate that about 10^{12} g of methane were released to the atmosphere by this practice in this region in 1973. Considering the looser production practices in other parts of the world, for example, the Persian Gulf, underwater venting may be contributing several teragrams (10^{12} g = 1 Tg) of methane per year to the fossil component of the atmospheric methane source budget.

The second source of petroleum production-related methane is that photochemically produced by sunlight on asphalts. As much as 678,000 barrels per day of asphalt were produced in the US in 1984⁴. This may be converted to 3.28×10^{13} g carbon per year. In the laboratory we can get about 15% conversion of similar materials to methane by thermal processes. Also we can show that the Sun's ultraviolet radiation can produce methane from asphalt. Using the estimates given

above, potentially 5 Tg of CH_4 per year could be formed from newly produced asphalt by photochemical and thermal induced processes. Considering worldwide asphalt production, its methane producing potential is even larger. Asphalt has a long lifetime in the environment. It is used extensively for road surfaces and roofing, both excellently exposed to the Sun's radiant energy.

We recommend that these sources of fossil methane be thoroughly evaluated in future studies of methane budgets.

WILLIAM M. SACKETT

TIMOTHY R. BARBER

Department of Marine Science,
University of South Florida,
Florida 33701, USA

1. Lowe, D. C., *et al.* *Nature* 332, 522–525 (1988).
2. Brooks, J. M., Bernard, B. B. & Sackett, W. M., in *Fate and Effects of Hydrocarbons in Marine Ecosystems and Organisms* (ed. Wolfe, D. A.) 373 (1977).
3. *Sea Technol.* 16, 10 (1974).
4. Cantrell, A. *Oil Gas J.* 32, 111 (1984).

Sexual selection in guppies defended

SIR—In a second criticism¹ of our finding² that predation risk determines female mating preferences in wild guppies, Houde restates Endler's points³, which we have previously answered⁴, and raises new objections. But these could be raised against most studies in evolutionary biology, which can at best include only a small number of populations and cannot positively preclude all alternative explanations.

First, Houde objects that we studied only two high-predation and two low-predation populations. But a strength of our investigation is that we examined more than one population, compared to previous studies of female preference, and we are presently examining more populations from a wider range of environmental conditions.

Second, Houde states that female preference must be studied under a wider range of experimental conditions, such as open versus partitioned aquaria. We agree, and in our earlier reply to Endler⁴ we suggested how our system using models of male guppies is ideally suited to provide such information.

Finally, Houde suggests that more information on selection in natural populations is necessary. In a companion article⁵ on variation in female preferences tested with live males, published in a format allowing a more expansive discussion, we have listed the types of data on selection that are needed for a full understanding of the evolution of female preference in this species. These include the determination of whether direct selection acts on females, whether the optimum male phenotype, the intensity of selection around the optimum, or both factors

change between high- and low-predation populations, and whether selection against brightly coloured male offspring is so great in high-predation populations that the linkage disequilibrium brought about by assortative mating by preferential females can never get established. We acknowledge limitations to our study, but reiterate that it provides strong evidence that selection against the attractive male character has led to correlated changes in female preference, as predicted by Fisher's runaway process of sexual selection.

FELIX BREDEN
GREGORY STONER

Division of Biological Sciences,
University of Missouri,
Missouri 65211, USA

1. Houde, A.E. *Nature* **333**, 711 (1988).
2. Breden, F. & Stoner, G. *Nature* **329**, 831–833 (1987).
3. Endler, J.A. *Nature* **332**, 593–594 (1988).
4. Breden, F. *Nature* **332**, 594 (1988).
5. Stoner, G. & Breden, F. *Behav. Ecol. Sociobiol.* **22**, 285–291 (1988).

Mycorrhizal infection and plant species diversity

SIR—Grime *et al.*¹ argue that the presence of mycorrhizas leads to an increase in the plant species diversity of laboratory microcosms. They grew species-rich mixtures of plants with and without infection by mycorrhizal fungi, and found a small but significant increase in plant species diversity as a result of mycorrhizal infection. The mechanism they invoke involves “the improved yields of the subordinate species [which] were not simply due to the reduced vigour or increased root/shoot biomass of the infected canopy dominant”. The small, suppressed plants are assumed to obtain carbohydrate from the larger, dominant plants via shared mycorrhizal hyphae. We take issue with this and propose a simpler explanation.

One criticism is that the radiolabelled CO₂ studies^{2,3} cited by Grime *et al.* as showing net carbohydrate flow into suppressed plants, in fact show carbohydrate movement only between plants. We agree there is clear evidence of carbon transfer through the mycorrhizal network, but not that there is a net flow of carbon to suppressed ‘sink’ plants. The obvious experiment of labelling CO₂ in the subordinate plants has never been done.

Grime *et al.* suggest that the “export of assimilate from ‘source’ (canopy dominant) to ‘sink’ (understorey components) through a common mycorrhizal network may be an important element of the mechanism maintaining species-rich communities on infertile soils.” We propose a simpler interpretation. Although ‘diversity’ increased in the mycorrhizal plots, species richness was identical under the two treatments; the increase in diversity results from a reduction in dominance following mycorrhizal infection. The dominant

plant in both treatments (*Festuca ovina*) declined in shoot weight by more than 30% following infection, and clearly loses more than it gains from mycorrhizal association. A substantial part of the response in the diversity index can be attributed to this alone (of the total increase in diversity of 47%, over half is accounted for by reduced dominance alone).

We suggest that the improved growth of previously suppressed species can be attributed to competitor release following mycorrhizal infection, coupled with the undisputed potential of mycorrhizae to enhance the uptake of phosphorus and nitrogen in nutrient-limited conditions. We see no need to invoke net carbohydrate flow into suppressed plants and would prefer to wait for direct evidence before accepting this more complicated explanation.

Furthermore, it is most unlikely that the result reported by Grime *et al.* is general; in this case mycorrhizal infection was detrimental to the growth of the dominant plant, but there will be others in which the dominant plant benefits from mycorrhizal association, and in becoming more dominant, brings about a reduction in plant species diversity.

J.M. BERGELSON

Department of Zoology NJ-15,
University of Washington,
Seattle, Washington 98195, USA

M.J. CRAWLEY

Department of Pure and Applied Biology,
Imperial College, Silwood Park,
Ascot, Berks SL5 7PY, UK

GRIME *ET AL.* REPLY—In grasslands on shallow calcareous soils, grazing has long been implicated as a control on plant species diversity, but the widespread occurrence of mycorrhizal infection in plants growing in such situations has been largely ignored. The objective of our experiment was to determine the impact of these two factors on the outcome of competitive interactions between species in microcosms. To examine effects arising from infection by mycorrhizal fungi, plant species known to differ in their susceptibility to infection were included, ranging from *Rumex acetosa* (not mycorrhizal) to *Centaurea erythraea* (highly infected). The year-long experiment indicated that infection promoted diversity and, in particular, increased the biomass of some of the subordinate species.

Bergelson and Crawley suggest that this treatment, like grazing, operated through selective debilitation of the dominant species, *Festuca ovina*. We agree that grazing and mycorrhizal infection are likely to have brought about competitive release, but Bergelson and Crawley seem to have overlooked features of the data which indicate effects peculiar to mycorrhizal infection, and not attributable to the reduced vigour of the dominant *F. ovina*.

In particular, the heavily mycorrhizal species *C. erythraea* showed a spectacular improvement in survivorship and yield when infected, but showed no evidence of competitive release when grazing reduced the yield of the dominant plant to levels considerably below those achieved in the mycorrhizal treatment. More telling still, inoculation with mycorrhizal fungi produced no effect on the yield of the non-mycorrhizal species *R. acetosa*.

We investigated possible contributory factors to the growth responses by feeding ¹⁴CO₂ to donor plants of *F. ovina*. Radioactivity levels were much higher in infected plants, providing, as we stated, support for earlier suggestions that export of assimilate through a common mycorrhizal network may be an important element of the mechanism maintaining diversity. We did not put the case more strongly because the experiment was not designed to determine net carbon movement. Earlier studies however, have shown that the magnitude of carbon transfer from potential donors to receivers can be influenced by shading of the receivers. This suggests that the polarity of movement through mycorrhizas is determined by source-sink relations and that the shaded subordinate plants in our experiment would receive more carbon than the well-illuminated dominants. The uninfected species *R. acetosa* supported this contention as it was the only one that failed to accumulate labelled carbon.

Furthermore, mycorrhizal infection is known to be a drain upon the carbon reserve of the host plant as carbon allocation to roots can increase by 10% on infection^{4,5}. Although such a drain is apparently sustained without adverse effect by dominant plants, it could be a problem for suppressed individuals, and it seems reasonable to consider that in such cases they are supplied with carbon by transfer from other individuals in the interconnected system.

We agree with Bergelson and Crawley that experiments to determine net flow of carbon to suppressed plants by labelling subordinate individuals are desirable. We have some simpler microcosms with single and two-species combinations which will enable such investigations to be made.

J. P. GRIME

J. M. MACKAY

S. H. HILLIER

D. J. READ

Unit of Comparative Plant Ecology
(NERC) and Department of Plant
Sciences,
The University, Sheffield S10 2TN, UK

1. Grime, J.P., Mackay, J.M.L., Hillier, S.H. & Read, D.J. *Nature* **328**, 420–422 (1987).
2. Francis, R. & Read, D.J. *Nature* **307**, 53–56 (1984).
3. Read, D.J., Francis, R. & Finlay, R.D. *Ecological Interactions in the Soil* (Blackwell Scientific, Oxford, 1985).
4. Pang, P.C. & Paul, E.A. *Can. J. Soil Sci.* **60**, 241 (1980).
5. Snellgrove, R.C., Splitsoesser, W.E., Stribley, D.P. & Tinker, P.B. *New Phytol.* **92**, 75–81 (1982).

Interpreters of dreams

J. Christian Gillin

The Dreaming Brain. By J. Allan Hobson. *Basic Books: 1988. Pp. 319. \$22.95.*

IF DREAMING, as Freud once said, "is the royal road to the unconscious", it has also long stimulated much of man's conscious speculation about human nature and experience, madness and creativity, and the sacred and prophetic. With the discovery of rapid eye movement (REM) sleep, by Aserinsky and Kleitman in 1953, however, dreaming became inextricably linked to fundamental physiological processes. Dreaming is an inherent part of normal REM sleep. Whether or not they say they normally dream, or are happy or sad, subjects who are awakened from REM sleep report dreams about 80 per cent of the time but only about 10 per cent from awakenings from the rest of sleep. We sleep about a third of our lives and spend nearly 20 to 25 per cent of that time in REM sleep, so we dream about 50,000 hours during a normal life span.

Over the past 30 years, much has certainly been learned about the phenomenology of dreaming and, perhaps, even more about the physiological basis of REM sleep. Nevertheless, these observations have not resulted in new theories about dreaming to displace older psychoanalytical views in the popular imagination or in the psychiatric clinic. Moreover, the usual void between mind and brain has rarely been breached in any serious exposition of the origins, content or form of dreaming.

With the publication of *The Dreaming Brain*, Allan Hobson presents an ambitious attempt to close that gap. Building on ideas first developed with his former colleague, Robert McCarley, he enlarges upon the 'activation-synthesis hypothesis' of dreaming, which is explicitly derived from premises about the neurophysiological basis of REM sleep. Both Hobson and McCarley are psychiatrists who have pursued electrophysiological, neuroanatomical and neuropharmacological studies of sleep in animals for over a quarter of a century. Their ideas about dreaming first arose, therefore, in the animal laboratory in searching for the physiological basis of REM sleep, rather than at the patient's bedside. Later, to test some of the predictions of the activation-synthesis hypothesis, they examined dream reports of their own, of subjects sleeping in the laboratory and, in this book, of an unidentified scientist who kept a detailed dream diary which Hobson secured.

According to the hypothesis, the experience of dreaming during REM sleep follows the *activation* of forebrain

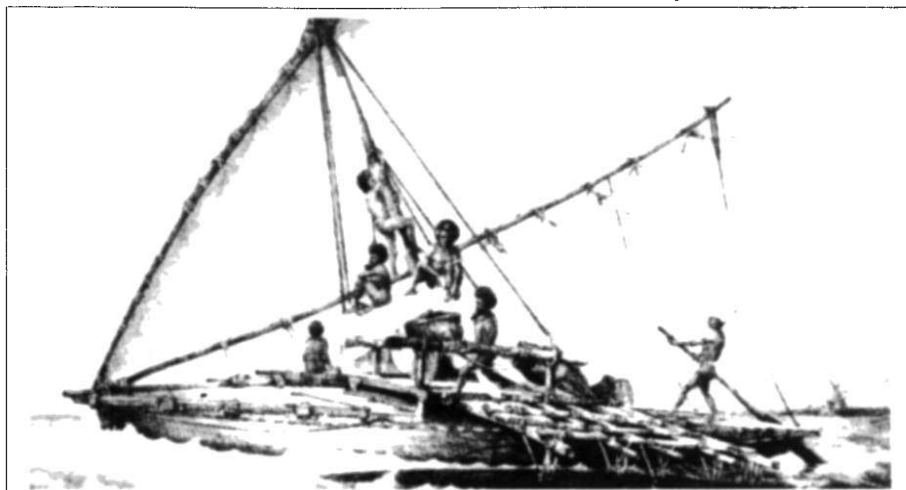
synthetic systems by arousal mechanisms ascending from within the brainstem. Non-specific 'sensory stimuli', such as pontine-geniculate-occipital (PGO) waves, are generated internally within the brainstem and bombard the forebrain during REM sleep. The patterns of these stimuli are often chaotic and the forebrain, which 'sees' these stimuli as if it were awake, attempts to 'make sense of them' by drawing upon memory stores and generating hallucinatory perceptions. Fundamental to this hypothesis is the concept of brain-mind isomorphism, the similarity of form between the physiological and psychological domains. For example, the subjective experience of moving during a dream is assumed to be related to patterned activation of motor systems and of those brain structures which orientate the person in space.

When Hobson and McCarley published the first outlines of the activation-synthesis hypothesis in the *American Journal of Psychiatry* in 1977, their paper attracted more letters to the editor than any single article ever published in that journal. Many of the letters were critical of the idea, primarily because of the two psychiatrists' overt scepticism of the psychoanalytical view of dreaming. The present book is unlikely to be more acceptable to the psychoanalysts because the author remains relatively uninterested in the 'deeper' personal meaning or background of dreaming. In Hobson's defence, however, he repeatedly states that the synthetic side of the hypothesis needs

more work — that is, it cannot yet "account for thematic narrative constancies of dreaming in any systematic way" (p. 271).

Neither is it certain that the thesis of the book will be accepted by neurophysiologists. The neurophysiology proposed to support the subjective experience of dreaming will strike some as naive, premature and untestable. For example, too much is known about the visual system to accept readily the idea that non-specific ascending stimulation, such as PGO waves, which are carried to the primary visual cortex, can be experienced as, for example, a beautiful landscape. Indeed, the book does not even address the issue of which visual pathways are required for the experience of vision or which motor systems are needed for the experience of movement during dreams. Studies of dream content and form in patients with specific neurological lesions might establish this important information.

Furthermore, there is the traditional 'chicken and egg' problem which so often attends discussions of the mind-brain. For example, Hobson correctly notes that single-cell neuronal activity is generally increased in the motor cortex during REM sleep. There is little doubt that this neural activity is related to motor movement which would be enacted were it not for the peripheral somatic muscle atonia normally occurring during REM sleep. For example, there are patients with a rare, recently described disorder called REM behaviour disorder, in whom muscle tone is preserved during REM sleep and who move around while dreaming. But what comes first, the mentation or the movement? Hobson would seem to argue that we dream out our acts rather than we act out our dreams. That is, because the neurons in the motor cortex or the vestibular complex fire rapidly during REM sleep as a result of the activation by brainstem, the fore-



Sailing by — a lithograph of a Fijian voyaging canoe, first published in Dumont d'Urville's Voyage au pôle Sud et dans L'Océanie (Paris, Gide, 1846). The picture is reproduced in Prehistory in the Pacific Islands by John Terrell, just published in paperback by Cambridge University Press, price £12.50, \$19.95.

brain must concoct a hallucinated subjective experience of movement which corresponds with the activation of the motor system. This argument harks back to the James-Lange theory of emotion: we are fearful because we tremble rather than we tremble because we are fearful.

Nevertheless, this is a provocative book which, in the long run, could be important, if only for its heuristic effect. Some 35 years after the discovery of REM sleep, Hobson is forcing us to come to grips with fundamental questions. He writes well and the book is nicely illustrated. To set the stage, he presents an interesting history of the parallel developments during the past century in the neurosciences, dream research and psychoanalysis. For these reasons, most of the ideas are readily accessible to the layman, but the specialist will also find much here that is new or interesting. This book will have served its purpose if it stimulates others to consider dreaming from a new psychobiological perspective. One day, perhaps, this will activate others to formulate a new synthesis incorporating both a comprehensive neurobiological and a psychosocial view of dreaming. □

J. Christian Gillin is a Professor in the Department of Psychiatry, University of California San Diego, and San Diego Veterans Administration Medical Center, La Jolla, California 92093, USA.

ADVERTISEMENT

The ANS Journals



Your number one source for current information about

NUCLEAR ENERGY



NUCLEAR NEWS

NUCLEAR SCIENCE & ENGINEERING

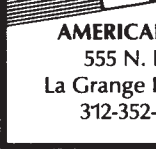


NUCLEAR TECHNOLOGY

TRANSACTIONS



REMOTE SYSTEMS TECHNOLOGY PROCEEDINGS



FUSION TECHNOLOGY

For free catalog contact:

AMERICAN NUCLEAR SOCIETY
555 N. Kensington Avenue
La Grange Park, Illinois 60525 USA
312-352-6611 Telex: 4972673

Reader Service No.22

A growing field

Clive Stace

Grass Systematics and Evolution. Edited by Thomas R. Soderstrom, Khidir W. Hilu, Christopher S. Campbell and Mary E. Barkworth. *Smithsonian Institution Press*: 1988. Pp.474. \$45, £35.

THE appearance of another sizeable book on grasses should come as no surprise, for the Poaceae, with about 10,000 species, are not only one of the biggest families of plants but by far the most important in terms of human welfare. Our civilizations built up around their cultivation, and our continuing success depends upon them no less today. In recent years several other large families have been the subject of symposia resulting in volumes that have immediately become standard texts. Those on the Apiaceae, Solanaceae, Fabaceae and Asteraceae are excellent examples. That we have had to wait so long for the present volume is a measure of the overwhelming significance of grasses and the vast number of publications that relate to them. On this basis, we should not have to wait too long for a sequel, as there are inevitably many gaps and biases in the book under review.

Until well into the twentieth century grass classification was patently artificial, being little but elementary pigeon-holing of taxa. This was due to the extreme reduction of floral parts and the stereotyping of an immensely successful vegetative structure, and it led to the deliberate search for cryptic micro-characters that might provide the basis for more natural classifications. That such characters were found in abundance, initially by Avdulov (cytology) and Prat (leaf epidermis) in the 1930s, is part of a success story which has resulted in the Poaceae having a more natural and hence valuable classification than virtually any of the other large families. Hence there are hardly any taxonomic novelties in the present work.

Grass Systematics and Evolution consists of 33 chapters based on papers presented at a symposium held at the Smithsonian Institution in July 1986. It covers topics as wide-ranging as structural diversity (from gametophytes to leaf epidermis), biochemical diversity (from chloroplast-DNA to isozymes and flavonoids), reproductive biology, and systematics of the principal groups (the five widely recognized subfamilies each have at least one chapter), including discussion of both phenetic and cladistic methods. As indicated in its title, topics such as plant breeding, ecology and relationships with man and other animals are omitted or scarcely covered. Some subjects that, based upon chapter titles, appear to be treated are found to receive only cursory attention.

For example, "Man and the Grasses: a History" (R. W. Pohl) consists of little over three pages of general chatter, and "Hybridization and Polyploidy in the Poaceae" (J. M. J. de Wet) can do scant justice to the subject in under six pages. Indeed, the vast body of data that has been accumulated under the general headings of cytogenetics and biosystematics is greatly under-represented. It is a lacuna that should be filled in a future volume, along with related topics such as plant breeding, including (as mentioned by Melvin Calvin in the foreword) genetic engineering.

Nevertheless, most of the chapters are written authoritatively, often in a review style, and provide not only an up-to-date picture of many aspects of agrostological research but a source for the basic literature (there is a general bibliography of 71 pages and in all over 2,300 references are included). Readers will frequently have to go to other publications for more details, but this is a book that must be considered the most valuable single work on the family. □

Clive Stace is a Professor in the Department of Botany, University of Leicester, Leicester LE1 7RH, UK.

Uniting nations

R. V. Short

World Population and the United Nations: Challenge and Response. By Stanley P. Johnson. *Cambridge University Press*: 1988. Pp.357. Hbk £37.50, \$59.50; pbk £12.95, \$22.95.

THIS is a Tio Pepe of a book, too dry for the palates of all but the most dedicated *aficionados* of the subject. The central theme is a retrospective analysis of the activities of the United Nations Fund for Population Activities (UNFPA), created by the United Nations in 1969 and directed until 1987 by that inspired internationalist, Rafael Salas. UNFPA has been responsible for raising large sums of money for population projects, but has tended to use other UN organizations such as UNESCO, ILO, FAO, UNDP, UNICEF and WHO as well as NGOs as its executing agencies. If you are already confused, then take heart — the book begins with a glossary of acronyms.

The text is replete with lengthy quotations from UN documents, speeches and minutes of meetings, backed up with extensive tabulations of data, and four graphs and a pie chart by way of light relief. The author is a professional administrator and sometime European parliamentarian with several books of fact and fiction to his credit. But although he

has chosen a subject of global significance and transcending importance to write about, the book seems curiously divorced from reality.

If we look at the spectacular achievements in population control in recent years — China, Thailand, Singapore — and the failures — India, Kenya, Egypt, to name but a few — can we honestly say that the United Nations Fund for Population Activities played a decisive role? There can be no denying that financial assistance from UNFPA has helped capitalize the successes, but it was probably lack of personal motivation and hence political will rather than shortage of funds that accounted for the failures.

One might have expected a book with such a challenging title to be replete with references to contraception, but the subject isn't even mentioned in the first 100 or so pages of preamble, and the outstanding achievements of WHO's Special Programme of Research, Development and Research Training in Human Reproduction are noted only in passing in a postscript.

With the wisdom of hindsight, could it be that politicians and administrators have got it all wrong? Dr Mahler, the Director-General of WHO for the past 15 years, and an outspoken champion of family planning, pointed out recently that the greatest health disparity in the world today is in pregnancy-related maternal mortality, which ranges from 1 in 10,000 in developed countries to 1 in 15 in some developing ones. To this we must also add the disparity in infant mortality, which ranges from less than 10 per 1,000 in developed countries to 250 per 1,000 or more in many areas of the developing world. Until we can begin to redress these reproductive health inequalities on a country-by-country basis, we can hardly expect individuals to pay much heed to political pronouncements about optimal rates of population growth. Whoever planned a family with global priorities uppermost in their mind? This is where the challenge lies and where the response is needed. It takes a barefoot doctor-turned-administrator, like Dr Mahler, to bring us all back to reality, and his impending retirement is a real setback.

Johnson's book will be a useful source of reference for those concerned with past developments in this field. We would surely all agree with him that population is a subject in which the world needs the co-ordinated action of all the relevant UN agencies. Although the United Nations behaves somewhat like a convoy of ships in wartime, forced to move at the pace of the slowest member-state, convoys get through in the end — if the Commodore knows where he is going. □

R. V. Short is a Professor in the Departments of Anatomy and Physiology, Monash University, Melbourne, Victoria 3168, Australia.

Conversations with a computer

L. Henry Shaffer

The Computer and the Mind: An Introduction to Cognitive Science. By P. N. Johnson-Laird. *Harvard University Press/Fontana*. 1988. Pp. 444. Hbk £23.50, \$29.50; pbk £6.95.

THE computer metaphor of mind has been in currency for some time. For anyone who wants to understand it better, Johnson-Laird's book is a very good starting point. Simply, a computer, stripped of considerations of hardware and architecture, is a concept within the mathematical theory of machines, and is defined by the abstract formalism of a Turing machine. Its main functions are to define computation, which is a systematic way of translating one symbol into another, and the limits of computability. Because symbols can be made to represent anything we like, any nonrandom process can be described as a computation. Thus the metaphor gives explicit meaning to the idea that mental processes are nonrandom.

It is a long way from supposing that the mind, or brain, is a computer to understanding how it achieves its results. Cognitive science has grown in the past decade as a multidisciplinary inquiry into the cognitive mind — how it can perceive the world, reason, solve problems, evolve and use language, plan action, and so on. Philosophy, linguistics, computer science, psychology and neuroscience have contributed their particular methods towards developing computational models of cognitive phenomena. For the full flavour of the research one has to visit laboratories around the world and see the little Frankenstein's that can almost speak, almost recognize patterns, almost understand language or almost pour a cup of tea. One can then appreciate the variety and ingenuity of the solutions achieved so far, and the enormous problems that remain to be solved before these laboratory artefacts come close to anything like human skill.

The book is presented as an introduction to cognitive science, but Johnson-Laird may have misjudged his aim in this. He tends to describe the idea of computation rather than computational modelling as such. His accounts of actual simulation studies often fail to convey the feel of solving problems. Instead, the computational method is illustrated by didactic exercises which are often too trivial to capture the spirit of the real thing. But this is a minor criticism. He writes on cognition with enviable clarity and wit, and with a breadth of vision that allows him to use music, art and literature in a natural way to make his points.

When he is on the home ground of his own research — language, reasoning and creativity — the writing is incisive and full of insights. Johnson-Laird's theoretical contribution to the concept of a mental model may not be formal enough to have explanatory power, but it seems intuitively correct and the concept is likely to remain central in cognitive theory for some time. Elsewhere he hits some dull patches: the section on learning, memory and action is rather desultory, and even connectionist machines, one of the most innovative areas in current cognitive research, seems not to have fired his imagination. The section on vision is good but still misses the sense of discovery conveyed, for instance, in David Marr's book of 1982. The last section, on computation and consciousness, contains entertaining but insubstantial explorations on the fringes of the computational metaphor.

How good is the computer metaphor of mind? There is no doubt that it has and can continue to throw light on aspects of human cognition. Its major test is in what it has to say about creativity. Johnson-Laird provides a lively discussion of how the chord sequences that form the basis of jazz improvisation may be created within a generative grammar. Here, of course, he stands on the shoulders of Noam Chomsky, who had the radical insight that a generative syntax allows the creation of an infinitude of linguistic forms. This kind of creativity is restricted to certain combinatorial problems, however, and we have to look elsewhere to understand how melodic invention or the flow of ideas in conversation come about.

Mathematical and scientific discoveries depend on finding significant analogies between phenomena and the properties of known structures. George Polya discussed how such innovative problem-solving may arise from heuristically controlled random raids on existing knowledge. This manner of using knowledge is pervasive even at mundane levels of everyday activity, and your Irishman in a pub spinning rhetoric by the yard provides as good an example as any. The point of these kinds of creation is that their goals confer a commitment but little constraint on the direction of the activity. They depend on an unbounded curiosity or a desire for social interaction. Such creativity has little to do with computation, but although he admits to doubts, Johnson-Laird seems unwilling to concede this point. Perhaps he would allow that such creativity has so far completely defeated attempts to include it in artificial intelligence. Anyone who thinks the computational metaphor of mind may be adequate in principle should consider a Sartrean hell in which one has to engage in eternal conversation with a computer. □

L. Henry Shaffer is a Senior Lecturer in the Department of Psychology, University of Exeter, Exeter EX4 4QJ, UK.

The spice of life

Anthony Hallam

Genetics, Paleontology, and Macroevolution. By Jeffrey Levinton. Cambridge University Press: 1988. Pp. 637. £27.50, \$37.50.

IN HIS famous concluding paragraph of *The Origin of Species* ("there is grandeur in this view of life"), Charles Darwin expressed wonder that the diversity of the organic world can have been produced from one or a few ancestors by the operation of several natural laws. For many if not most people today, the outstanding question about evolution remains the same as it did in Darwin's day — given descent from a common ancestor, how did the extraordinary diversity of life in space and time come about?

The dominant evolutionists for much of this century have been population geneticists, but it has become increasingly apparent that, although the preferential replication of genes by means of natural selection may well be a necessary condition for evolution to take place, it does not provide a sufficient answer to the fundamental question. Hence the current interest in macroevolution, which is widely understood as evolution above the species level, but which is defined more precisely by Levinton as the sum of those processes that explain the character-state transitions by which we diagnose the evolutionary differences between the main taxonomic ranks. Treatment of this vast subject requires coverage of a wide range of disciplines, namely systematics, genetics, molecular and developmental biology, functional morphology, ecology and palaeontology, a daunting task that the author undertakes in masterly style.

In Levinton's opinion, the single most important fact that palaeontology has contributed to evolutionary biology is that the history of life has witnessed a reduction of body-plan diversity and a stabilization of form. A large number of phyla occur first as fossils in a brief time-span near the beginning of the Cambrian. Subsequently, most of the organic turnover has occurred at lower taxonomic levels. In explanation, Levinton proposes a model whereby functional, developmental and genetic 'ratchets' have combined to ascribe a high burden to any prospective change. At lower taxonomic levels, long-term stability is also apparent. Thus many genera and families are remarkably homogeneous phenotypically. The fossil record

nevertheless amply demonstrates that phyletic change is common and sufficient to generate large-scale evolutionary change. But although morphological stasis is apparent in many fossil lineages, Levinton ascribes it primarily to the action of stabilizing selection, and strongly opposes the punctuated equilibrium model of evolution whereby virtually all evolution takes place at speciation events.

The clustering of organisms into discrete morphological groups, often with big gaps between them, is considered to be the consequence of a combination of factors (divergent evolution, extinction of intermediates, and functional and constructional limitations). The study of mass extinctions has demonstrated that many organisms disappear from the face of the Earth as a result of chance happenings rather than inferior adaptations, such that extinction events can exert a sort of downward causation in opposition to the kind of thing beloved of the more reductionist biologists. On the other hand,

recent research in molecular genetics offers great promise in at last getting to grips properly with the relationship between genotype and phenotype, hitherto the most intractable of subjects.

There is much to admire in this book. Levinton writes well and demonstrates an impressive grasp of a large number of more or less subtle concepts. In his trenchant defence of the neodarwinist position, what he has to say is for the most part well argued, balanced and reasonable. Unfortunately, in his determination to dismiss the arguments of the leading proponents of punctuated equilibrium, he tends to be too extreme in his criticisms. Levinton gives insufficient credit to those such as Eldredge, Stanley and Gould who have challenged the conventional wisdom and have thereby brought palaeobiology to the forefront of evolutionary studies, where it is likely to remain. □

Anthony Hallam is Lapworth Professor of Geology in the University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Common sensors

Joseph Jordan

Chemical Sensors. Edited by T. E. Edmonds. Blackie, Glasgow/Chapman & Hall, New York: 1988. Pp. 326. £45, \$127

THE title, *Chemical Sensors*, reminds me of an episode that occurred 17 years ago when I received a telephone call from a prestigious funding agency inviting me to lecture at the "first workshop ever devoted to bioelectrochemistry". I accepted and asked, "but what is bioelectrochemistry?" Taken aback, the caller replied that my own research was known to be a prime example of bioelectrochemistry.

In a similar vein, I am tempted to inquire "What is a chemical sensor?". The answer seems to be that it is what each of the 19 contributors to this volume perceive it to be. Their perceptions vary, while that of the editor himself is somewhat contradictory. In the preface, Edmonds implies (correctly) that the pH glass electrode, which is the oldest and best-known ion selective electrode (ISE), is one of the most successful chemical sensors. Nine chapters later, the same Edmonds (and coauthor Birch) deplore as misguided the idea that ISEs qualify as chemical sensors at all. Notwithstanding such exercises in hair-splitting, this is an excellent book because all 14 chapters are devoted to timely topics, are well written and make good reading.

P. D. Beer's contribution, unilluminatingly entitled "Molecular and Ionic Recognition by Chemical Methods", is a masterpiece. It includes a superb summary of the work by Pedersen on crown ethers,

of Lehn on cryptands and of Cram on spherands, for which they shared the 1987 Nobel Prize in chemistry. This coverage either reflects prophetic insight or a remarkable coincidence: it is apparent from the copyright page that the book went to press before the announcement of the award.

The two chapters on voltammetric, amperometric and potentiometric transducers are commendably brief and competently written. However, they duplicate material treated in many contemporary textbooks and monographs. On the other hand, the opening contribution on molecular and ionic recognition by biological systems is a rather indiscriminate potpourri, ranging from antibodies and enzymes to hormones and DNA, all crammed into 12 pages. Not surprisingly, this chapter lacks depth and contributes little to the central theme of the book.

The editor's overall philosophy is succinctly stated in the preface: "This book is intended to be a handbook, giving practical information to its readers, as well as supplying ideas for future use". In that respect, the chapters dealing with MOSFETs, thick-film and catalytic devices, and spectroscopic, fibre-optic and piezoelectric transducers, impressed me as the most valuable parts of the book; indeed, the 63-page subsection on "Nonelectrochemical Transduction" is a unique feature. To my knowledge, there is no other book currently available which provides a bird's eye view of both electrochemical and non-electrochemical sensors. As such, it will be widely read. □

Joseph Jordan is a Professor in the Department of Chemistry, The Pennsylvania State University, 152 Davey Laboratory, University Park, Pennsylvania 16802, USA.

● **Erratum:** In the review of *Barnacle Biology* (*Nature* 333, 716; 1988) the book was said to be a *Festschrift* for J. H. Crisp. The Crisp concerned was of course Professor D. J. Crisp, and the volume was not a *Festschrift* as such but was dedicated to Professor Crisp by the contributors after it had been compiled.

Fertilizer risks in the developing countries

Gordon R. Conway and Jules N. Pretty

In industrialized nations nitrogen fertilizers are known to present a health hazard, and their use may be restricted. The developing countries are a different case altogether.

NITROGEN fertilizers demonstrably increase the yields of crops, and are essential to food production in many parts of the world. Yet, under certain circumstances, they pollute the environment and may contribute to fatal human illnesses¹. In some quarters, the evidence for such hazards in the industrialized countries is felt to justify restricting fertilizer use. Stringent regulations are being enforced on nitrate levels in drinking water: standards of 45 mg nitrate l⁻¹ (equivalent to 10mg NO₃-N) have been set by the US Public Health Service and the World Health Organization, while the European Economic Community has specified a limit of 50 mg nitrate l⁻¹. In Iowa the legislature has gone so far as to introduce a groundwater protection bill that would impose a tax on all nitrogen fertilizers.

The question is beginning to be asked: should similar restrictions be imposed in the developing countries? Pressure from environmental lobbies, as in the case of pesticide use, can rapidly result in effective limitations putting food production in jeopardy.

Average fertilizer rates to arable land are much lower in developing than in industrialized countries. Regional averages are about 30 kg N ha⁻¹ for Asia, 15 kg for Latin America and 4 kg for Africa, compared with averages of 188 kg for Western Europe and 146 kg for Japan². Nonetheless most of the recent increase in fertilizer use has occurred in developing countries (Fig. 1). Much of this growth has been associated with the introduction of 'Green Revolution' cereal varieties, which produce high yields but are dependent upon heavy applications of fertilizer. In some of the intensive wheat- and rice-growing regions of the developing countries average application rates are now approaching those of the industrialized countries.

Relative effects

Although it is difficult to separate the relative effects of fertilizers and other factors, studies of Asian rice production indicate that all fertilizers have contributed to about 24% of the growth in output since the mid-1960s, the remainder being mostly due to new varieties, irrigation and capital investment³. The combined consequences have certainly been dramatic: per capita food production in the develop-

ing countries has risen by 7% since 1964, and in Asia and Latin America, the regions most affected by the Green Revolution, by 27% and 9% respectively⁴.

The population of the developing countries is expected to grow by three billion by the year 2025, implying a need

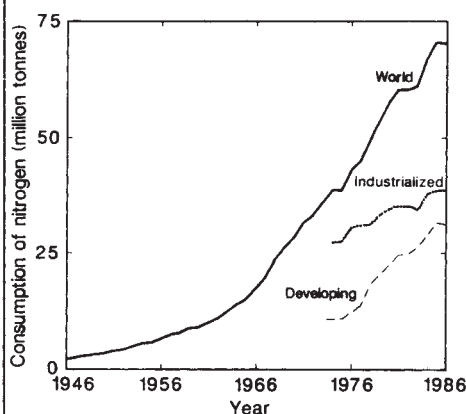


Fig. 1 Nitrogen fertilizer consumption 1946–1986 (ref. 2). In 1986 nitrogen comprised some 55% of global nitrogen + phosphorus + potassium consumption.

for a comparable growth in food production — or even greater growth, if rising expectations are to be met and food is to be provided for the many millions who are currently malnourished. In recent years arable cropping has expanded onto increasingly marginal lands and the potential for large-scale irrigation schemes is diminishing. Further growth in crop production, whether for local food consumption or for export, will have to come from intensification of agriculture, which currently means increasing fertilizer use.

Health hazards from the use of nitrogen fertilizers arise primarily because of the high losses of nitrogen to the environment, either as gases, or as nitrates which leach or drain to ground or surface waters and find their way into drinking supplies. Nitrates themselves are benign, but can be converted to hazardous nitrites.

The best understood hazard is methaemoglobinaemia. Nitrates are converted in the body by bacteria to nitrites, which then bind to haemoglobin, so impairing the capacity of the blood to transport oxygen. Most susceptible are infants during the first few months of life, exhibiting in extreme cases the blue-baby syndrome.

Methaemoglobinaemia is a rare condi-

tion. Most recorded cases in industrialized countries occurred during the 1940s and 1950s, in localities that relied on poorly constructed private wells which were contaminated by organic wastes from septic tanks and livestock. The conclusion from these cases is that methaemoglobinaemia usually results from high nitrate levels (over 100 mg l⁻¹) in drinking water that is also contaminated with bacteria and is used to make up dried-milk preparations for infants. A lack of dietary vitamin C, which enhances reduction of methaemoglobin, is a further contributory factor.

Namibia

For these reasons, the risk of methaemoglobinaemia may be considerably higher in developing countries. Yet we are aware of only one study of its occurrence, conducted in a rural area of Namibia⁵. Out of nearly 500 infants under one year old, about 8% had blood levels of more than 5% methaemoglobin, sufficient to cause the first clinical symptoms of oxygen starvation. The population was almost entirely dependent on ground wells for drinking water and 40% of the infants regularly consumed water at concentrations greater than 90 mg nitrate l⁻¹. The source of the contamination, though, was probably livestock waste rather than fertilizers.

By contrast, the link between nitrates and cancer is still contentious. The starting point is again the bacterial reduction of nitrate to nitrite, but this is followed by the combination of the nitrate with secondary amines and amides, derived from food and various environmental contaminants, to produce N-nitroso compounds. These are powerful carcinogens in a wide variety of animal species. Nitrosation has been shown to occur in human digestive tracts but it is still not known whether this results in human cancer.

The epidemiological evidence of a link between gastric cancer and nitrate levels in drinking water is conflicting¹. Recent studies have found negative correlations, and it seems that, although high nitrate levels in the diet and drinking water may play a role, they are not the main factor in determining the incidence of gastric cancer⁶⁻⁸. Despite the increase in nitrogen fertilizer use in the industrialized countries, gastric cancer mortality has fallen on average by 30% since the early 1950s. In

the developing countries, mortality rates from gastric cancer are comparable to those in the industrialized countries only in Chile, Colombia, Costa Rica and parts of Brazil, Shanghai and among the Chinese population of Singapore⁹. In these places, too, however, the rates are mostly falling.

In Chile, the only country in the world with natural nitrate deposits, a correlation was found between gastric cancer and nitrogen fertilizer use¹⁰, but there was no correlation with water nitrate levels¹¹. Urinary nitrate levels in children were found to be highest in the regions of low cancer risk, and while vegetables tended to have high nitrate levels in both high- and low-risk areas, the highest levels occurred in those grown in the low-risk areas¹².

One of the highest incidence rates of gastric cancer occurs in the Department of Narino in Colombia (150 per 100,000). In the early 1970s, a precursor stage of the cancer, chronic atrophic gastritis, was found in more than 50% of adults over 40 years of age¹³. However, less than 20% of the drinking water wells in the region contained more than 50 mg nitrate l⁻¹ and the source of nitrate is not known. Although the area is predominantly agricultural, commercial fertilizers were not being used intensively. People from the high-risk area had significantly higher salivary and urinary nitrate than those from low risk regions, though elevated urinary nitrate was also evident in those who did not drink well water¹⁴.

Bladder cancer

N-nitroso compounds may also be a causative factor in bladder cancer. In the Nile Delta, where the incidence is high and strongly correlated with infection by the parasite *Schistosoma*, there is also a high prevalence of infection of the bladder with bacteria capable of both reducing nitrate and catalysing nitrosation¹⁵. Although there is no information on nitrate contamination of drinking water, annual fertilizer applications are of the order of 300 kg N ha⁻¹. Urinary nitrate levels tend to be high, and in relatively healthy young men the highest concentrations of *N*-nitroso compounds occurred in those who were both infected with reducing bacteria and suffering from schistosomiasis¹⁵. Patients already showing clinical signs of bladder cancer also tend to have high levels of urinary *N*-nitroso compounds¹⁶.

Apart from these three, equivocal, situations there is no other evidence from the developing countries of links between fertilizer use, nitrate levels in drinking water and cancer, despite the occurrence in some places of very high levels of nitrate contamination of surface and ground waters.

According to various surveys in India

and Africa, 20–50% of wells contain nitrate levels greater than 50 mg l⁻¹ and, in some cases, as high as several hundred milligrams per litre¹⁷. But apart from surface waters in the vicinity of fertilizer factories there is little recorded evidence of high nitrate levels resulting from fertilizer use. In the developing countries it is usually wells in villages or close to towns that contain the highest levels, suggesting that domestic excreta are the main source, though livestock wastes are particularly important where drinking troughs are situated close to wells, often producing levels in the hundreds of milligrams per litre. Similarly, although many tropical rivers and lakes are commonly eutrophicated, the critical source of nutrient inflow is usually organic effluents.

Thus at present there is little evidence of health hazards arising from fertilizers in the developing countries. But will the picture change as usage increases? The potential hazards are probably greatest in those countries with pronounced seasonal climates where nitrates are flushed into surface and ground water at the onset of the rainy season. Such regions are also the main sites of large-scale irrigation. Some of the highest cereal yields in the world are produced by irrigation of crops in the dry season; these environments are already subject to heavy applications of fertilizer, which seem likely only to increase in the future. Irrigation water provides a direct conduit for nitrate to surface and ground water, and during re-use nitrates may become progressively concentrated.

Nitrogen losses, whether to the atmosphere or by leaching or surface drainage, may also be greater in the tropics. Dryland tropical crops lose up to 40–50% of applied nitrogen, but losses are seldom less than 60–70% for rice paddies^{18,19}. The amount lost is greatly influenced by the mode of application and the type of fertilizer. Urea fertilizer, for example, which accounts for over 70% of nitrogen consumption in Asia and is favoured for increased use, is subject to high losses of ammonia and nitrous oxide^{20,21}, gases that contribute to environmental acidification and global warming.

But, for the immediate future, health hazards in developing countries are more a function of poorer nutrition and hygiene. Diets are often deficient in vitamin C which appears to protect against the development of gastric cancer as well as methaemoglobinemia. Diarrhoea causes 5–10 million deaths annually, although it is not known how many of the cases result from pathogens capable of reducing nitrate or catalysing nitrosation. On the positive side, nitrate does not become concentrated in human milk²², and breast feeding appears to confer some resistance to both diarrhoea and the development of methaemoglobinemia. It is the use of dried-milk products requiring preparation

with water that is potentially hazardous.

In our opinion, none of the evidence we have discussed here or presented elsewhere¹⁷ justifies the restriction of fertilizer use in the developing countries. Nevertheless, it would be wrong to be complacent. Methaemoglobinemia may be more prevalent than is supposed, its presence masked by far greater incidence of diarrhoea. Investigation of the effects of the very high levels of nitrate in surface and ground water, that occur in some localities as a result of domestic and livestock wastes, could provide evidence of the possible effects of much higher usage of nitrogen fertilizer.

Restriction

If, in future, public pressure and scientific evidence combine to produce a worldwide restriction on fertilizer use, the developing countries will need to ensure that they are not unfairly disadvantaged by having their average fertilizer rates frozen at levels well below those of the industrialized countries. In the meantime, greater effort needs to be devoted to the promotion of more efficient and less polluting use of nitrogen fertilizers and to alternative biological sources of nitrogen. □

1. US National Research Council, *The Health Effects of Nitrate, Nitrite and N-nitroso Compounds* (National Academy Press, Washington, 1981).
2. *Fertiliser Yearbook* (Food and Agriculture Organization, Rome, 1986, 1985, 1984, 1977, 1976, 1973, 1967, 1961, 1955, 1953, 1951).
3. Herdt, R.W. & Capule, C. *Adoption, Spread and Production Impacts of Modern Rice Varieties in Asia* (International Rice Research Institute, Los Banos, 1983).
4. *Production Yearbook* (Food and Agriculture Organization, Rome, 1987, 1984, 1976, 1973).
5. Super, M. *et al. Wat. Res.* **15**, 1265–1270 (1981).
6. Forman, D., Al-Dabbagh, S. & Doll, R. *Nature* **313**, 620–625 (1985).
7. Forman, D., Al-Dabbagh, S. and Doll, R. *Nature* **315**, 462 (1985).
8. Beresford, S.A.A. *Int. J. Epidemiol.* **14**, 57–63 (1985).
9. Parkin, D.M. (ed.) *Cancer Incidence in Developing Countries* (IARC Scientific Publications, Lyon, 1986).
10. Armijo, R. & Coulson, A.H. *Int. J. Epidemiol.* **4**, 301–309 (1975).
11. Zaldivar, R. & Wetterstrand, W.H. *Z. Krebsforsch.* **92**, 227–234 (1978).
12. Armijo, R. *et al. Int. J. Epidemiol.* **10**, 57–62 (1981).
13. Cuello, C. *et al. J. natn. Cancer Inst.* **57**, 1015–1020 (1976).
14. Hawksworth, G. *et al. in N-Nitroso Compounds in the Environment* (eds Bogovski, P. & Walker, E.A.) 229–234 (IARC Scientific Publications, Lyon, 1975).
15. Hicks, R.M. *et al. Trans. R. Soc. trop. Med. Hyg.* **76**, 519–527 (1982).
16. El-Merzabani, M.M., El-Aaser, A.A. & Zakhary, N.I. *Eur. J. Cancer* **15**, 287–291 (1979).
17. Conway, G.R. & Pretty, J.N. *The Risks of Fertiliser Use in the Developing Countries: A Review* (International Institute for Environment and Development, London, 1988).
18. Greenwood, P.J. *et al. J. agric. Sci., Camb.* **95**, 471 (1980).
19. De Datta, S.K. *Principles and Practices of Rice Production* (Wiley, New York, 1981).
20. Fillery, I.R.P. & Vlek, P.L.G. *Fertil. Res.* **9**, 79–98 (1986).
21. Bolle, H.-J., Seiler, W. & Bolin, B. in *The Greenhouse Effect, Climatic Change and Ecosystems* SCOPE No. 29 (eds Bolin, B. *et al.*) 157–203 (Wiley, New York, 1986).
22. Green, L.C., Tannenbaum, S.R. & Fox, J.G. *New Engl. J. Med.* **306**, 1367–1368 (1982).

We thank David Baldock, Nigel Bell, Dame Barbara Clayton, Sir Richard Doll, David Forman and David Morley for their comments. The review was supported by the Ford Foundation.

Gordon R. Conway and Jules N. Pretty are at the International Institute for Environment and Development, 3 Endsleigh Street, London WC1H 0DD, UK. Gordon Conway is on leave of absence from Imperial College of Science and Technology, London.

Origin of the Hubble constant controversy

R. Brent Tully

Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

The long-standing argument as to whether the Hubble constant is closer to 50 or 100 km s⁻¹ Mpc⁻¹ may be due to large local disturbances in galactic motions caused by excess mass in the Coma-Sculptor cloud. Outside this region, a value of 85 to 95 km s⁻¹ Mpc⁻¹ is consistently found.

THE proposed downward revision of the Hubble constant by Sandage and Tammann^{1,2} to 50–55 km s⁻¹ Mpc⁻¹ sparked a debate that has continued for over a decade. A value in the range 90–100 km s⁻¹ Mpc⁻¹ has been advocated by others^{3,4}. Uncertainties have routinely been assessed at ~10%, so the conflicting results have appeared beyond reconciliation. I argue here that the dispute has its origins in a perversity of nature, referred to as the 'local velocity anomaly'. This anomaly misled Sandage and Tammann in two respects: first, it led them to a low value of the Hubble constant because their early work was heavily weighted by galaxies drawn from the region of the anomaly, and second, the anomaly creates the signature of Malmquist bias and caused Sandage and Tammann to overestimate the importance of this systematic effect.

Malmquist bias

Before turning to the matter of the local velocity anomaly, the point will be made that it is possible to measure bias-free distances to individual galaxies in the field. A relatively simple issue has become clouded by the discussion in the literature. To be specific, suppose distances are to be estimated using the luminosity–line-width relation⁵. The usual procedure is to use either the double-regression relationship or the single regression with uncertainties in luminosity. As Giraud⁶, Bottinelli *et al.*⁷ and Feast⁸ have shown, there certainly are Malmquist biases to contend with in these cases. Lynden-Bell *et al.*⁹ have an elaborate discussion on how to compensate for the bias in the context of their diameter–dispersion relationship for early-type systems.

Malmquist biases arise from sample magnitude cutoffs. In the cases mentioned above, there is a tendency to select galaxies brighter than what would be the mean in a complete sample. The distance to a galaxy that is brighter than the mean at a given line width will be underestimated if assigned the intrinsic luminosity of the mean. While the regression on luminosity is an example of a relationship that is likely to produce the standard Malmquist bias, it is instructive to define relationships that produce an anti-Malmquist bias. In this discussion we are assuming for simplicity that the luminosity–line-width relationship can be adequately described by a power law, that is, a linear relationship on a log-log plot.

This situation is illustrated in Fig. 1, which actually displays the material used to calibrate our operational relationship as described at the end of this section. The dashed line with the flattest slope represents the regression with uncertainties in luminosities that, if used to derive distances to field galaxies, produces estimates that suffer from Malmquist bias. Now consider the dashed line with the steepest slope that has no physical meaning, but was produced simply to induce an anti-Malmquist bias. Recall that the Malmquist bias arises because field galaxies tend to be drawn from above the assumed relationship, especially more distant galaxies that are drawn from the luminous end of the sample. But one can arbitrarily assume a relationship so steep that it becomes difficult to draw field galaxies from above the relationship at the bright end. Most luminous galaxies would lie below a very steep relationship,

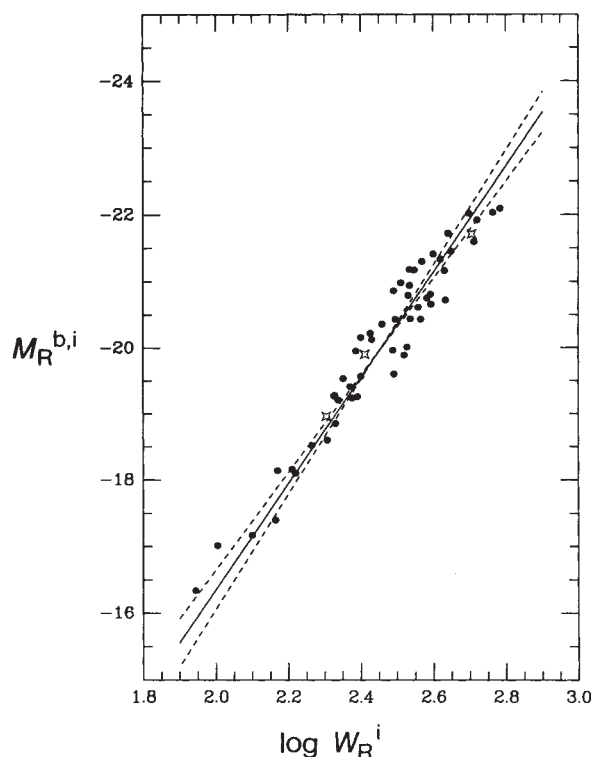


Fig. 1 Luminosity–line-width relation for the combined Virgo and Ursa Major clusters. The shallow dashed line is derived from the regression with uncertainties in luminosity. The steep dashed line represents an arbitrary relation that would generate a negative Malmquist bias. The solid line is derived from the regression with uncertainties in line width and provides the bias-free estimator. The absolute magnitude of the relationship is established by fitting to the three local calibrator systems: M31, M33 and NGC 2403 (open stars).

hence the anti-bias. The point is that the amplitude and even the sign of the Malmquist bias depends on the slope of the assumed relationship. Hence, there must be a slope that causes the bias to vanish. The next step is to examine a recipe that prescribes the slope needed for unbiased measurements.

Suppose that a galaxy in the field is selected based on its apparent magnitude, and there is no bias to the sample from H I profile line width (actual samples may not be free of line-width bias, but radio telescopes are now sufficiently sensitive for suitable samples to be obtainable). The bias is avoided if distances are determined based on the relationship that bisects the line-width distribution at any given luminosity. Given a complete sample, and assuming that there is a power-law relationship between the observables and a normal distribution about the mean log line width at a given luminosity, then the

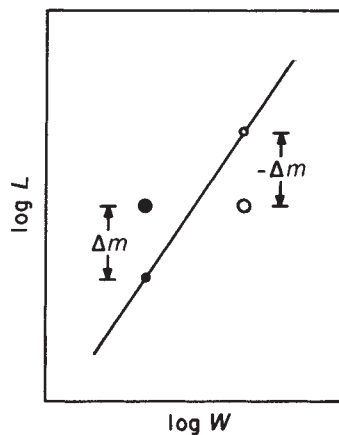


Fig. 2 Symbolic luminosity-line-width relation. If the straight line is derived from the regression with uncertainties in line width, then there is equal probability of drawing either the large closed circle or the large open circle from a magnitude-limited sample. The errors associated with drawing either of these objects are equal and opposite in sign.

desired relationship is provided by the regression with uncertainties in line width.

This idea can be repl...ed with recourse to Fig. 2 for clarity. Suppose that a field galaxy were observed that would really lie at the location of the large filled circle in Fig. 2. In attempting to derive its distance, its line width would be known, and it would be taken that it lies on the assumed relationship at the location of the small filled circle. Hence, the intrinsic magnitude of the galaxy would be taken to be too faint by an amount Δm . But the associated erroneous distance estimate is not contributing to a systematic bias if the assumed relationship bisects the intrinsic line-width distribution. In this case, it would have been equally probable to have observed the system with the same intrinsic luminosity indicated by the large open circle in Fig. 2, which would have led to an equal error of opposite sign in the distance modulus. The essential point is that as one samples to a given horizontal cut in the relationship illustrated in Fig. 2, corresponding to some apparent magnitude limit, there will be symmetry in the way objects are sampled from both sides of the intrinsic relationship and, hence, there will not be any systematic distance estimate bias.

This point has been made by several authors¹⁰⁻¹³, but the procedure has been obscured by complicated model fitting. Simply given the regression with uncertainties in line width determined from a suitable complete sample, then distances can be determined without bias to individual galaxies. The fit to the complicated inverted relationship discussed by Schechter¹⁰ and the straightforward prediction of a luminosity from a line width are geometrically equivalent. The key point is that the fit is being done to the regression on line widths.

Frankly, this proposal seemed too simple to be true, so some artificial data were generated to test it out in the specific context of the luminosity-line-width relationship. Four random numbers were generated. (1) A line width was randomly selected. (2) The line width implied a mean intrinsic luminosity according to a specification of the luminosity-line-width relation. The actual intrinsic luminosity assigned is taken to be this expected luminosity plus a randomly chosen luminosity difference drawn from a Gaussian distribution with a specified dispersion. (3) A random number was compared with the probability that the assigned luminosity would be drawn from a Schechter differential luminosity distribution of a specified form¹⁴. If the random number exceeds the value of the probability function, then this artificial galaxy is rejected from consideration. (4) If the galaxy survived, a distance was randomly assigned in a way that assured homogeneity with distance, with an outer cutoff.

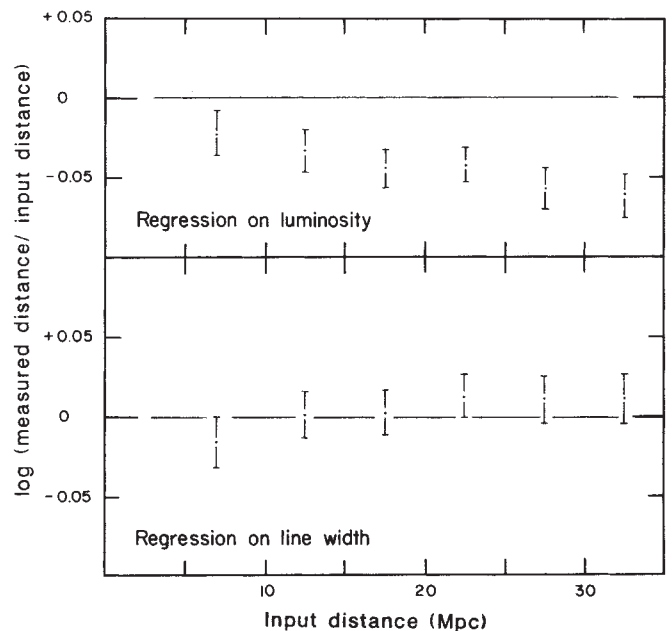


Fig. 3 Systematic trends in measured distance. Means of the ratio of measured to input distances for simulated data are plotted as a function of input distance for two methods of measurement. A point corresponding to a properly measured distance will lie on the horizontal line at zero on the ordinate. Means and standard deviations of the mean are given for groups of roughly 25 data points each, binned by input distance. In the top panel, distances are measured using the regression on luminosity to a complete sample. In the bottom panel, distances are measured using the regression on line width to the same complete sample. For the specific test that is illustrated, the luminosity-line-width relationship has an intrinsic luminosity dispersion of 0.50 mag.

For each test, a sample of a thousand galaxies was generated with properties that resembled the real data to be described later. The two complementary linear regressions were performed to the complete data sets. All galaxies were then rejected that were fainter than a specified apparent magnitude limit, and distances were determined for those that remained based on the separate regression relationships. Distances so determined could be compared with input distances and biases could be evaluated. The parameter Δ = distance modulus (input) - distance modulus (measured) was calculated, and the means for each test are gathered in Table 1. Figure 3 shows the dependence of Δ on input distance for the average of two selected tests. It is seen that there is always a bias in the case of regression on magnitudes. The bias is worse if the dispersion of the luminosity-line-width relation is large and the bias increases with input distance. There is modest bias if the double regression is used and no significant bias if the regression on line width is used.

Again, the key is a good formulation of the regression on line widths. Pierce and Tully¹⁵ have recently recalibrated the luminosity-line-width relations in B, R, I and H passbands using

Table 1 The dependence of Malmquist bias on regression procedure

	Input distance modulus - measured distance modulus		
	$\sigma = 0.25$	$\sigma = 0.50$	$\sigma = 0.75$
Regression on L	0.09 ± 0.03	0.22 ± 0.04	0.39 ± 0.06
Double regression	0.06	0.10	0.14
Regression on W	0.02	-0.02	-0.11

The tests were repeated with three different values for the intrinsic dispersion of the relationship between luminosity (L) and line width (W). The dispersions (σ) are given in the luminosity dimension, in magnitudes.

extensive data in the Virgo and Ursa Major clusters. They have samples that are essentially complete to magnitude limits that are faint compared with the luminosities of galaxies in field samples. The assumption is made that galaxies in the field are like the galaxies in these clusters. This assumption is probably a good one because the Ursa Major cluster, in particular, displays conditions very much like those in the field (cluster dispersion: 148 km s^{-1} ; crossing time: $0.5/H_0$; overwhelmingly a population of late-type, gas-rich systems). The R-band relationship for the current cluster sample is displayed in Fig. 1.

Galaxies lie in clouds

The author has recently described how >99% of nearby galaxies lie within entities called clouds¹⁶. Our galaxy resides in the Coma-Sculptor cloud. Most galaxies within 10 Mpc lie in either the Coma-Sculptor cloud or an appendage called the Leo Spur. The cloud assignments are based on kinematic distances. The interest now is to compare distances determined from the regression on line width to kinematically determined distances for all possible galaxies within a single cloud. For the present discussion, the H-band sample of Aaronson *et al.*¹⁰ will be used with the calibration given by Pierce and Tully¹⁵. Kinematic distances are alternatively based on the assumption of pure Hubble expansion, or on the Virgocentric retardation model discussed by Tully and Shaya^{17,18}.

For both alternative cases, the ratio R_m/R_k (measured distance/kinematic distance) was determined for individual galaxies and then the mean of this ratio within separate clouds was found. On the average, the ratio should be unity, across all the clouds if kinematic distances are being calculated based on the proper value of the Hubble constant. Significant departures would indicate some clouds have motions not described by the kinematic model.

The two panels of Fig. 4 display the ratio R_m/R_k as a function of mean measured distance for those clouds with at least five independent distance estimates. Figure 4a illustrates results with the pure expansion case and Fig. 4b demonstrates the situation under the assumption of the specific Virgocentric retardation model. There are two especially important points to emphasize. (1) In *a*, the case of the pure expansion model, a value of $80 \text{ km s}^{-1} \text{ Mpc}^{-1}$ was required to force the mean of the ratio of observed to predicted distance to unity; in *b*, with the specific Virgocentric mass model assumed, a much higher value of $H_0 = 95 \text{ km s}^{-1} \text{ Mpc}^{-1}$ was required (in this case, after three- σ rejection of three deviant points). (2) With both kinematic models, though particularly in the Virgocentric perturbation case, the points associated with the Coma-Sculptor cloud in which we live and an appendage to our cloud called the Leo Spur are well above the mean.

The first matter has a natural explanation. In general, an observer who resides within a region influenced by a mass concentration and who is trying to determine the Hubble constant from measurements confined to the same region will tend to find a value for H_0 that is biased too low. It has long been understood that H_0 should be measured on large scales to minimize the uncertainties caused by peculiar velocities. But to this author's knowledge, it has not been appreciated as a generalized phenomenon that if peculiar velocities arise from gravitational perturbations then observers close to the perturbation will tend to underestimate H_0 . In our case, the model that presumes the larger gravitational perturbation for consistency requires the larger value of H_0 .

The second matter probably has the same explanation. For the moment, accept the proposition that the Virgocentric retardation model provides a better description of the Local Supercluster environment than the pure expansion model. Although from Fig. 4b the model seems adequate for most clouds, it badly fails to describe distances in our cloud and the contiguous Leo Spur. Within these entities, both of the kinematic models proposed underestimate distances substantially.

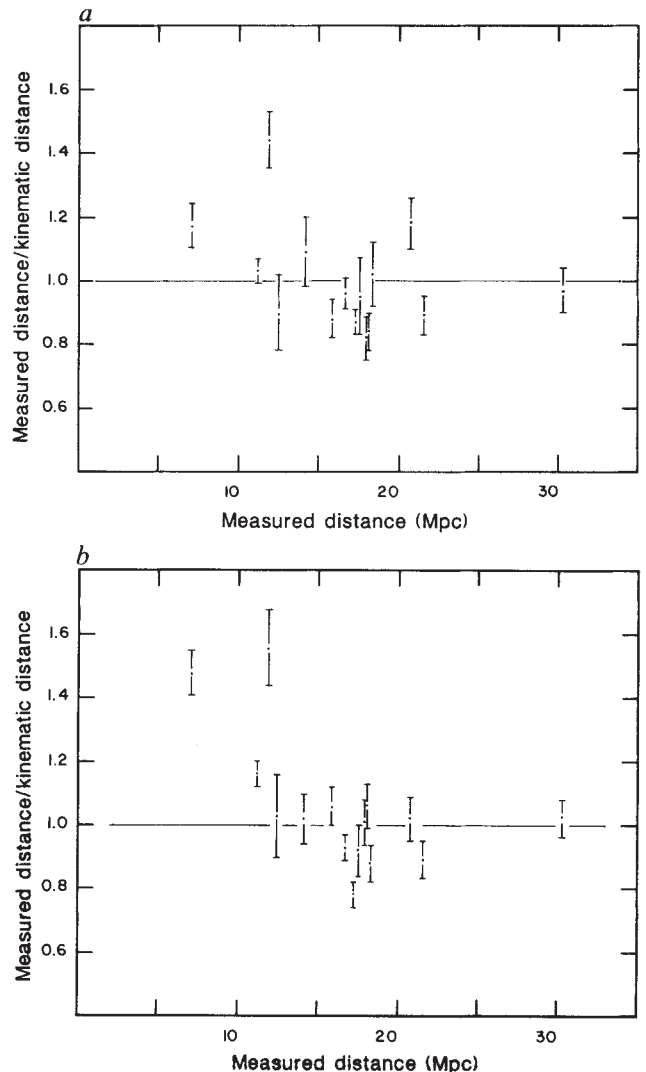


Fig. 4 *a*, Ratio of measured to expected distances: pure expansion model. Each data point corresponds to the mean ratio associated with all galaxies with measured distances within a single cloud. Only clouds with at least five measurements are included. Model assumes $H_0 = 80 \text{ km s}^{-1} \text{ Mpc}^{-1}$. *b*, Ratio of measured to expected distances: Virgocentric retardation model. Data points correspond to the same clouds as in *a*. Model assumes $H_0 = 95 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Figure 5 illustrates values of H_0 that would be derived from each cloud separately. Open circles follow from the pure expansion model and filled circles follow from the Virgocentric flow model. If only galaxies within our cloud and its spur are considered then, with either kinematic model, one would determine $H_0 \approx 63 \text{ km s}^{-1} \text{ Mpc}^{-1}$ with the zero-point to the distance estimator scale used here. As already pointed out, the general effect of the Virgocentric perturbation model is to elevate H_0 . It is this enhancement of the difference between the two nearby clouds and the more distant ones that causes the two nearby clouds to stand out particularly in Fig. 4b.

The effect seen in the local cloud and its spur will be termed the 'local velocity anomaly'. Velocities associated with galaxies in this region are lower than expected from either free expansion or any sort of spherical Virgocentric flow model. The most natural explanation is a repetition of the phenomenon already suspected to exist on the scale of the Local Supercluster. There is probably retardation of motions caused by mass associated with the Coma-Sculptor cloud and Leo Spur.

The phenomenon of H_0 increasing with distance as seen in Fig. 5 has been noted before (for example, see refs 7 and 19)

and attributed to Malmquist bias. Alternatively, Malmquist bias would act to decrease R_m/R_k as a function of distance. If it were believed that Malmquist bias is the dominant effect and that the nearest clouds are least affected, then $H_0 \approx 63 \text{ km s}^{-1} \text{ Mpc}^{-1}$ would be expected. But it should be clear that Malmquist bias would not produce the abrupt step, then plateau with distance, that is observed in Figs 4a or 4b. If the Coma-Sculptor cloud and Leo Spur are ignored, the data are consistent with the proposition of the previous section that Malmquist bias has been avoided.

Kraan-Korteweg, Cameron, and Tammann¹³ have identified a phenomenon related to the local velocity anomaly, but were unable to explain it and concluded they were probably seeing an observational artefact. De Vaucouleurs²⁰ actually found a manifestation of the local velocity anomaly three decades ago, which he interpreted in the context of rotation of the Local Supercluster.

The mass of the Coma-Sculptor cloud

The obvious interpretation of the local velocity anomaly is that expansion of the local cloud is being retarded by self-gravity. The Coma-Sculptor cloud and the Leo Spur must be falling together. The retardation velocities are modest, but they are still a substantial fraction of the expansion velocities of galaxies that are so nearby. A more detailed analysis is in preparation, but an order-of-magnitude estimate of the mass that must be present can be obtained from the equation used by Lynden-Bell *et al.*⁹ to determine the mass of the 'great attractor'

$$\delta M \sim RV|\Delta V|/G$$

Characteristic parameters are $V = 450 \text{ km s}^{-1}$, $\Delta V = -200 \text{ km s}^{-1}$, and $R = 7.1 \text{ Mpc}$, whence the excess mass $\delta M = 1.3 \times 10^{14} h^{-1} M_\odot$ ($h = H_0/100$).

The luminosity associated with the galaxies in the Coma-Sculptor cloud is well determined (assuming not much is hidden by the Milky Way). Extrapolating mildly based on an assumed luminosity function¹⁴ and correcting for absorption, $L_B = 4.3 \times 10^{11} h^{-2} L_\odot$. Hence, the mass-to-light ratio is $300h M_\odot/L_\odot$. This value, characteristic of a region about 10 Mpc across, is roughly twice the value of $125h M_\odot/L_\odot$, which is characteristic of halos on a scale of $0.3h^{-1} \text{ Mpc}$ around galaxies¹⁶, but much smaller than the closure density value of $1,600h M_\odot/L_\odot$.

Summary

Most galaxies within 800 km s^{-1} lie in either the Coma-Sculptor cloud or the Leo Spur. Galaxies that are members of these clouds have anomalously low velocities for their distances. This claim is made based on distance estimates that should be free of Malmquist bias. The local velocity anomaly cannot be described by any reasonable spherical Virgocentric retardation model.

Historically, galaxies within the two nearest clouds have contributed heavily to low estimates of H_0 (ref. 1), and the apparent increase of H_0 with distance has been attributed to Malmquist bias^{7,19}. This phenomenon has probably also contributed to the speculation by Segal²¹ that there is a quadratic relationship between distance and velocity. As we live in an overdense region, it is not surprising that a local measurement of H_0 leads to an underestimate.

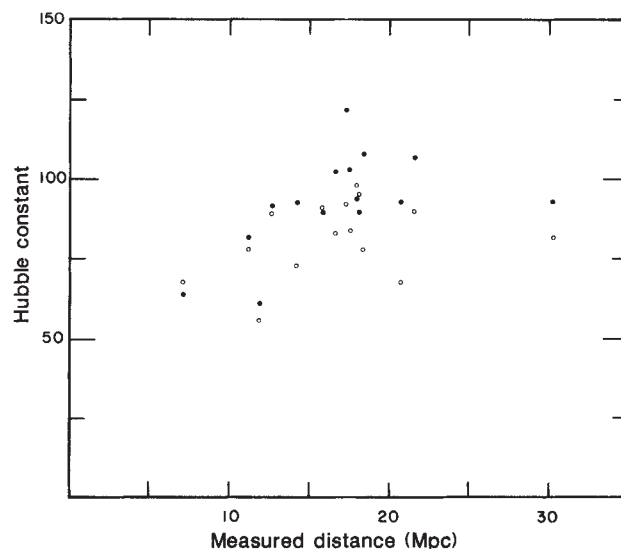


Fig. 5 Hubble constant derived from individual clouds. The value of H_0 required to force the mean ratio of observed to expected distances to unity is given for each of the clouds of galaxies that contribute to Fig. 4. Open circles: pure expansion model; closed circles: Virgocentric retardation model.

Intercomparison of the results between the two kinematic models considered shows that the preferred value of the Hubble constant is sensitive to the choice of model. There is some evidence that the Virgocentric mass model tried is somewhat excessive because, whereas in the pure expansion case the ratio of measured to kinematic distances for clouds in the opposite hemisphere from the Virgo cluster tend to lie below the mean, they tend to lie above the mean with the mass model. Hence with the zero-point used¹⁵, H_0 is expected to be greater than the value of $80 \text{ km s}^{-1} \text{ Mpc}^{-1}$ required by the free expansion model, but is probably less than the value of $95 \text{ km s}^{-1} \text{ Mpc}^{-1}$ determined by the specific Virgocentric mass model tested here.

Aaronson *et al.*⁴ have determined H_0 from observations of clusters out to $10,000 \text{ km s}^{-1}$, hence over a scale larger than the scales of known peculiar velocities. They found $H_0 = 88 \text{ km s}^{-1} \text{ Mpc}^{-1}$ on the same zero-point used here. This determination of the Hubble constant is within the range of the present estimates and is to be preferred, given the model dependency of the locally derived values.

The most straightforward explanation of the local velocity anomaly posits that there is enough mass within the Coma-Sculptor cloud to cause retardation from Hubble expansion within this cloud and to attract the Leo Spur. Roughly $1.4 \times 10^{14} M_\odot$ is required, whence $M/L \approx 270 M_\odot/L_\odot$ on a scale of 10 Mpc if $H_0 = 90 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

More detailed modelling is underway in collaboration with Ed Shaya and Mike Pierce, and I thank them for contributions that have influenced the present discussion. This research has been supported by a grant from the US National Science Foundation.

Received 25 January; accepted 10 May 1988.

1. Sandage, A. & Tammann, G. A. *Astrophys. J.* **196**, 313–328 (1975).
2. Sandage, A. & Tammann, G. A. *Nature* **307**, 326–329 (1984).
3. de Vaucouleurs, G. in *Galaxy Distances and Deviations from Universal Expansion* (eds Madore, B. F. & Tully, R. B.) 1–6 (Reidel, Dordrecht, 1986).
4. Aaronson, M. *et al. Astrophys. J.* **302**, 536–563 (1986).
5. Tully, R. B. & Fisher, J. R. *Astr. Astrophys.* **54**, 661–673 (1977).
6. Giraud, E. *Astr. Astrophys.* **174**, 23–27 (1987).
7. Bottinelli, L., Gouguenheim, L., Paturel, G. & Teerikorpi, P. *Astr. Astrophys.* **156**, 157–171 (1986).
8. Feast, M. W. *Observatory* **107**, 185–188 (1987).
9. Lynden-Bell, D. *et al. Astrophys. J.* **326**, 19–49 (1988).
10. Schechter, P. L. *Astr. J.* **85**, 801–811 (1980).

11. Aaronson, M., Huchra, J. P., Mould, J. R., Schechter, P. L. & Tully, R. B. *Astrophys. J.* **258**, 64–76 (1982).
12. Teerikorpi, P. *Astr. Astrophys.* **141**, 407–410 (1984).
13. Kraan-Korteweg, R. C., Cameron, L. & Tammann, G. A. in *Galaxy Distances and Deviations from Universal Expansion* (eds Madore, B. F. & Tully, R. B.) 65–72 (Reidel, Dordrecht, 1986).
14. Tully, R. B. *Astr. J.* **95** (July, 1988).
15. Pierce, M. J. & Tully, R. B. *Astrophys. J.* **330** (July 15, 1988).
16. Tully, R. B. *Astrophys. J.* **321**, 280–304 (1987).
17. Tully, R. B. & Shaya, E. J. *Astrophys. J.* **281**, 31–55 (1984).
18. Tully, R. B. *Nearby Galaxies Catalog* (Cambridge University Press, Cambridge, 1988).
19. Sandage, A. *Astrophys. J.* **331** (August 15, 1988).
20. de Vaucouleurs, G. *Astr. J.* **63**, 253–266 (1958).
21. Segal, I. E. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2473–2477 (1975).

Spliceosomal RNA U6 is remarkably conserved from yeast to mammals

David A. Brow & Christine Guthrie

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

The small nuclear RNA U6 and its gene have been isolated from yeast. In striking contrast to other yeast spliceosomal RNAs, U6 is very similar in size, sequence and structure to its mammalian homologue. The single-copy gene is essential. These properties suggest a central role in pre-mRNA processing. An extensive base-pairing interaction with U4 snRNA is described; the destabilization of the U4/U6 complex seen during splicing thus requires a large conformational change.

THE removal of intervening sequences from pre-messenger RNA is an essential part of gene expression in eukaryotes. In mammals, it requires the participation of a number of factors, including at least five small nuclear RNAs (snRNAs) designated U1, U2, U4, U5 and U6 (see ref. 1 for a review). The RNAs are present as individual ribonucleoprotein complexes (snRNPs), with U4 and U6 being base-paired together in a single particle^{2,3}. These snRNPs assemble onto the pre-mRNA substrate to generate the active spliceosome. The U1 and U2 snRNPs interact directly with specific sites on the pre-mRNA, and there is evidence that the U5 snRNP does so also. The role of the U4/U6 snRNP in splicing, however, is unknown.

Pre-mRNA splicing in the yeast *Saccharomyces cerevisiae* proceeds by the same general mechanism used in mammals. It was therefore expected that homologues of at least some of the mammalian U RNAs would be present in yeast. In fact, the yeast and mammalian snRNA repertoires have diverged to the point that most of the yeast U RNAs were identified only after extensive characterization of the large family of yeast small RNAs⁴⁻⁹. The structure of the first yeast U RNA to be discovered was quite surprising: *S. cerevisiae* U2 RNA is more than six times as long as its human counterpart (1,175 instead of 189 nucleotides), and almost all the sequence similarity is in the first 80 nucleotides, where the two RNAs are 75% identical⁵. Likewise, yeast U5 (ref. 6) and U1 (refs 7, 8) RNAs are respectively 1.5 and 3.5 times the length of their mammalian counterparts with which they share only limited sequence and structural similarity. Recently we identified the yeast U4/U6 RNA complex⁹ and found that, in contrast to the other spliceosomal RNAs, U4 and U6 are very close in size to their mammalian counterparts. The primary structure of yeast and human U4 RNA is significantly different, but yeast U6 is sufficiently conserved in sequence to be readily detectable by hybridization with a mouse U6 gene⁹.

Here we report the isolation and characterization of the yeast U6 RNA (snR6) and its gene (*SNR6*). U6 turns out to be the most highly conserved spliceosomal snRNA: snR6 (112 nucleotides) is 75% identical to human U6 (106 nucleotides) over more than half its length. The yeast U6 gene is essential and has two upstream sequence elements that are also found upstream of mouse and human U6 genes. Sequence comparisons lead us to propose a conserved secondary structure of the U4/U6 interaction domain. The yeast U4/U6 RNA complex has a melting temperature of 53 °C, more than 25 °C above the temperature at which splicing is carried out *in vitro*. Spliceosomes assembled *in vitro* and isolated on native gels lose U4 but not U6 before the pre-mRNA is cleaved^{10,11}. Our results suggest that an active mechanism is likely to be employed to achieve this destabilization of the U4/U6 complex.

Yeast U6 gene structure

We first identified snR6 as a component of an RNA complex that was isolated on native gels and contained a previously

characterized snRNA, snR14. snR14 is 160 nucleotides long and shares some sequence similarity with mammalian U4 (145 nucleotides)⁹. snR6 was isolated from the complex and used for partial enzymatic sequencing (see Fig. 1 methods) and as a probe to isolate a genomic clone of its gene, *SNR6*. A 1.8-kilobase-pair (kb) subclone containing the gene and approximately equal amounts of 5' and 3' flanking DNA was constructed (Fig. 1a) and the sequence of 500 base pairs (bp) encompassing the gene was determined (Fig. 1b). The RNA sequencing revealed that the 3' end of snR6 is heterogeneous and contains 4, 5, 6 or 7 contiguous U residues. The 5' end of snR6, identified by primer extension of total yeast RNA, is homogeneous. Thus yeast U6 comprises a set of RNAs 112–115 nucleotides in length (Fig. 1b). Comparison of the *SNR6* coding region with mammalian U6 RNA reveals about 60% identity over the length of the two sequences, increasing to more than 80% identity in the middle third of the RNAs (see below).

In animals, U6 RNA appears to be synthesized by RNA polymerase III (pol III)¹²⁻¹⁴ although the promoter includes a TATA box¹⁴⁻¹⁶ and an octamer motif¹⁴⁻¹⁷, recognition signals normally associated with genes transcribed by RNA polymerase II. The heterogeneity of the 3' end of yeast snR6, and the fact that it falls in a long oligo(dT) stretch (Fig. 1b) are consistent with snR6 being a primary transcript made by yeast pol III¹⁸. Furthermore, the snR6 gene is transcribed in a yeast cell-free pol III transcription system (D.A.B. and C.G., unpublished observations). We find no consensus pol III intragenic promoter elements in the yeast U6 gene, which is not surprising in view of the fact that vertebrate U6 genes containing deletions of a putative 'A box' element¹⁵, or even the entire coding region¹⁴, were found to be efficiently transcribed.

Upstream of the snR6 gene are two regions strikingly similar in sequence to elements implicated in initiation of transcription of the mammalian U6 gene. One is centred 30 nucleotides 5' of the presumptive snR6 initiation site and contains a consensus TATA box (Fig. 1c) with a perfect 10-bp match between the mouse and yeast genes. Although the conserved U6 TATA box has not been shown to be an essential promoter element, a triple point mutation in the TATA box upstream of another gene transcribed by pol III, the human 7SK RNA gene, greatly reduces its transcription *in vitro*¹⁹. Further upstream, between positions -60 and -40 of the yeast and human U6 genes, is an 11 nucleotide perfect match (Fig. 1c). Deletion analysis of the human U6 promoter maps the 5' border of an essential element to this region, between -67 and -43 (ref. 16). Although the U6 -30 and -50 regions are less well conserved in other animals^{20,21}, the high degree of similarity between yeast and mammals strongly suggests they are part of the *SNR6* promoter.

Yeast U6 is essential

To determine whether U6 performs an essential function in yeast, we generated a diploid heterozygous for a disrupted allele of *SNR6* and assessed the fate of its haploid spores. The yeast

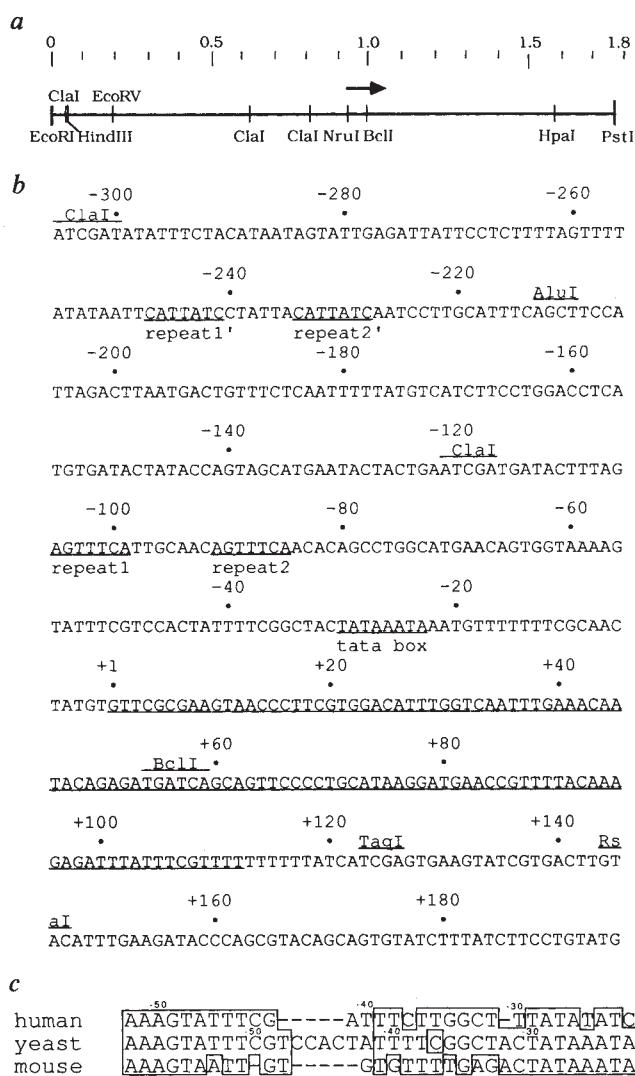


Fig. 1 **a**, Restriction map of the *EcoRI*-*PstI* fragment containing the yeast U6 RNA gene, *SNR6*. This fragment was cloned into pUC118 to generate pEP6. The arrow above the map shows the position and direction of transcription of *SNR6*. The scale indicates the distance in kb from the *EcoRI* site. There is a *dam* methylated *ClaI* site 0.1–0.5 kb from the *EcoRI* site. Enzymes that did not cut are: *AccI*, *BamHI*, *BglII*, *BstEII*, *NcoI*, *SacI*, *SmaI*, *StuI*, *XbaI*. **b**, Nucleotide sequence of the non-template strand of *SNR6*. The sequence corresponding to the shortest transcript (112 nucleotides, see text) is underlined. A 'TATA box' and two pairs of upstream direct repeats are indicated, as are relevant restriction sites. **c**, Alignment of sequences upstream of yeast, human¹⁶ and mouse^{14,31} U6 RNA genes (see text). Numbers indicate distance in nucleotides from the presumptive transcription initiation site. **Methods.** 1.3×10^{10} cell equivalents of yeast splicing-extract Fraction I, prepared as described²⁸ from *S. cerevisiae* IH2004 (a/ α , *leu2*⁺, *trp1*/*trp1*, *ura3*-52/⁺, *prb1*-1122/*prb1*-1122, *pep4*-3/*pep4*-3, *prc1*-407/*prc1*-407, *his1*⁺, *mal*⁺, *gal2*⁺, obtained from A. Mitchell and I. Herskowitz), was combined with an equal volume of 100 mM TrisCl, pH 7.4, 200 mM NaCl, 4% SDS, and proteinase K (E. Merck) was added to a final concentration of 2 mg ml⁻¹. After 2 h at 20 °C, the mixture was extracted three times at room temperature with 1:1 phenol/CHCl₃ and once with CHCl₃, then precipitated with ethanol. RNA was end-labelled by ligation to [5'-³²P]pCp as described³², except that DMSO was omitted and incubation was at 0 °C for 24 h. The mixture was diluted with an equal volume of 15 mM EDTA, 100 mM NaCl, 15% glycerol and subjected to two-dimensional 'RNP gel' electrophoresis³³. *snR6* was located by autoradiography (see Fig. 6 of ref. 9), excised from the gel and eluted at 37 °C in 0.3 M Na acetate, 0.1 mM EDTA, 0.1% SDS. After precipitating with ethanol, *snR6* was again end-labelled with [5'-³²P]pCp, but in 10% DMSO at 4 °C, yielding 3.7×10^5 c.p.m. (Cherenkov). One third of this RNA was used to probe a partial *Sau3a* yeast genomic library in YEp13. From the initial 15-kb insert isolated, a 1.8-kb *EcoRI*-*PstI* fragment was subcloned into pUC118 (pUC18 with nucleotides 5,465–5,941 of M13 inserted at the *NdeI* site) to generate pEP6. The remainder of the *snR6* was further purified on a 6% polyacrylamide, 8.3 M urea gel and the shortest species was subjected to partial enzymatic sequencing^{34,35}. Oligonucleotide 6A, complementary to positions 43–56 of *snR6*, was synthesized and used to map the 5' end of *snR6* by primer extension of total yeast RNA³⁶. The gene was sequenced by the dideoxy method using 6A and other gene-specific oligonucleotides, as well as M13 universal primers with subclones of *SNR6*.

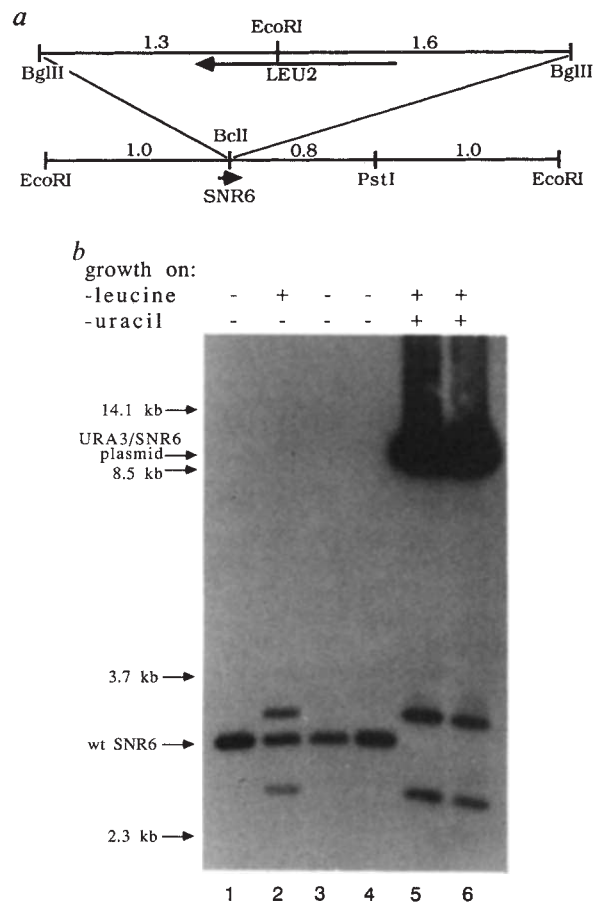
LEU2 gene was inserted into the *BclI* site in the middle of *SNR6* (Fig. 2a), and this disrupted allele was introduced into a *leu2*⁻ *ura3*⁻ diploid by one-step gene replacement²². The resulting *Leu*⁺ transformants contained both the wild-type and disrupted *SNR6* alleles (Fig. 2b, lane 2). When a *Leu*⁺ transformant was sporulated, only two spores from each of 11 tetrads were viable. Viable spores were always *Leu*⁻, indicating that they had inherited the non-disrupted *SNR6* allele. In contrast, when the *SNR6* heterozygote was first transformed with a *URA3*-marked centromere plasmid into which the 1.8-kb *snR6* gene-containing fragment was inserted, tetrads with two, three or four viable spores were obtained (in roughly equal proportion). In this case, *Leu*⁺ spores were obtained and such spores were always *Ura*⁺ as well, indicating the presence of the plasmid-borne *SNR6* allele. Furthermore, *Leu*⁺ haploids did not grow on medium containing 5-fluoro-orotic acid (5-FOA). This uracil analogue is lethal to cells carrying the wild-type *URA3* allele²³, so lack of growth indicates that the *SNR6/URA3* plasmid is required for viability. In contrast, *Leu*⁻ *Ura*⁺ spores always grew on 5-FOA plates, and thus were able to lose the plasmid. Southern analysis of genomic DNA from each spore of a four-spore tetrad confirmed that the *SNR6/URA3* plasmid co-segregated with the *SNR6::LEU2* allele (Fig. 2b, lanes 3–6). We conclude that *SNR6* is an essential, single-copy gene, and that it is adequately expressed when present on a low-copy-number plasmid along with 0.9 kb of 5' and 0.7 kb of 3'-flanking DNA.

U6 conservation

Comparison of the *snR6* sequence with published U6 sequences yields insights into the structure of U6 and the region of U4 RNA with which it interacts. Figure 3a displays an alignment of the yeast, bean, fly and mammalian U6 sequences. On the basis of conserved structural features, we have divided U6 into four domains, which we will address in turn. The 5' terminal domain is highly variable in length, and is responsible for almost all the difference in size of the U6 RNAs. Potential base-pairing sequences in the mammalian RNAs suggested that this region forms a stem/loop structure^{24,25}, and fly U6 contains a single, compensatory base-pair change in the putative stem²⁰. The yeast U6 can form an extended version of this stem, with three additional base pairs (Fig. 3c). Surprisingly, although *snR6* contains the sequence predicted to be in the loop in animals (5'UUCG3'), in yeast this falls within the predicted stem. Apparently the upper stem/loop can accommodate substantial variation in length and sequence. In contrast, the base of the stem is conserved in sequence as well as in structure: the first, second and fourth base pairs are invariant, as is the dinucleotide 3' of the stem (Fig. 3c), which delimits the 5'-terminal domain. In bean U6, the 5' half of the terminal stem appears to be missing (Fig. 3a)—but this, and the fact that pol III very rarely initiates transcription with a pyrimidine, suggests that the bean U6 sequence may be incomplete at its 5' end. The recently deter-

Fig. 2 *a*, Restriction map of the *LEU2* disruption of *SNR6*. A *Bgl*II fragment containing the yeast *LEU2* gene was inserted into a unique *Bcl*I site in pEP6 at position 56 of the *snR6* gene. Insertion of the construct into the chromosome by homologous recombination at the *SNR6* locus is expected to generate the *SNR6::LEU2* allele shown. Numbers indicate the approximate distance in kb between restriction sites. Arrows indicate the position and direction of transcription of genes. *b*, Southern analysis of *Eco*RI-digested genomic DNA from *SNR6::LEU2* transformants. DNA from the diploid parental strain (lane 1), a *Leu*⁺ transformant (lane 2), or all four spores of a tetrad derived from a *Leu*⁺ transformant carrying wildtype *SNR6* on a *CEN4/ARS1/URA3* plasmid (lanes 3–6) was probed with ³²P-labelled pEP6. The ability (+) or inability (–) of each isolate to grow on media lacking leucine or uracil is shown above the lane. Approximate expected sizes of hybridizing fragments are: *SNR6*, 2.8 kb; *SNR6::LEU2*, 2.3 kb and 3.4 kb; *SNR6* plasmid, 9.8 kb. The rescue plasmid contains vector sequences present in the probe and so labels very efficiently.

Methods. The 2.9-kb *Bgl*II fragment of YEp13 (ref. 37) was ligated to *Bcl*I-cut, phosphatased pEP6 prepared from *dam*[–] *Escherichia coli* to generate pEP6::LEU2. This plasmid was digested with *Hind*III and *Pst*I (see Fig. 1*a*) and the mixture was used to transform *S. cerevisiae* IH1788 (*a/α*, *trp1/trp1*, *leu2/leu2*, *ura3/ura3*, *his4/his4*, *can1/can1*; obtained from A. Mitchell and I. Herskowitz) by the Li acetate method³⁸. A *Leu*⁺ diploid was further transformed with YCpEP6 (provided by K. Shannon), which consists of the 1.8-kb *Eco*RI–*Pst*I fragment of pEP6 ligated to *Bam*HI linkers and inserted into the *Bam*HI site of YCp50 (ref. 39). A *Leu*⁺*Ura*⁺ diploid was sporulated and genomic DNA was prepared from each spore of a tetrad, as well as from a *Leu*⁺*Ura*[–] diploid and the untransformed parent, digested with *Eco*RI, electrophoresed on a 0.7% agarose gel and transferred to nylon membrane (Hybond-N). The blot was probed with 50 ng (15 × 10⁶ Cherenkov c.p.m.) pEP6 labelled with ³²P by nick translation.



mined *Caenorhabditis elegans* U6 sequence (J. Thomas and T. Blumenthal, personal communication) and that from the yeast *Kluyveromyces lactis* (H. Roiha and C.G., unpublished results) fit the consensus shown in Fig. 3*c* perfectly, and extend the range of variability in stem/loop length (14 to 30 nucleotides).

Sequence conservation falls off for a short distance downstream of the 5'-terminal domain, but then increases dramatically before the U4:U6 interaction domain (discussed below) is reached (Fig. 3*a*). In this central domain is a region that, in mammals, is predicted to base pair with a region in the 3' terminal domain to form a stem of substantial stability^{24,25}. No similar stem can be formed in *snR6*, however, so this is not a conserved structural feature. We can suggest no alternative conformation for the central domain, but the invariance of its length, and the high sequence conservation of its 3' half, imply an essential function. The 3' terminal domain is highly divergent in sequence, although almost constant in length. A stem/loop proposed for this region² of mammalian U6 cannot be formed in yeast, and we are unable to identify alternate structures.

U4/U6 interaction

The U4/U6 interaction domain is comprised of two regions that have been proposed to base pair with U4 (stems I and II, Fig. 3*a*). Pairing of the 5' end of U4 with a region of U6 (stem II) was proposed² on the basis of sequence complementarity in the mammalian RNAs. Fly and bean U6 RNAs are identical to mammalian U6 in this region and hence provide no opportunity to look for compensatory (base-pair maintaining) changes in U4. Yeast U6, however, differs at five positions in this stretch. Three of these clearly exhibit compensatory base changes in yeast U4 (ref. 9) (Fig. 3*d*, positions 7, 8 and 16 in *snR14*), providing strong phylogenetic evidence for the existence of the proposed stem. The 5'-adjacent segment of human U6 has been crosslinked with psoralen to a complementary region in U4 (ref.

26) (Fig. 3*a*, stem I) and this stem can also be formed with compensatory changes in yeast (Fig. 3*d*). The formation of these two stems leaves a large portion of U4 looped out (*snR14* positions 19–56 (ref. 9)). We find that this region can form a stem that is perfectly conserved between yeast and mammals with respect to its length, the position of a bulged nucleotide and the identity of its terminal base pairs. This conservation is remarkable in view of the fact that all 13 nucleotides comprising the stem's central portion change identity between yeast and mammals (Fig. 3*d*). The loop constrained by this stem exhibits strong sequence conservation.

The predicted free energy of formation of the U4/U6 complex via base-pairing of stems I and II is very similar for the yeast and mammalian RNAs. If stems I and II are treated as a single continuous double helix, the calculated ΔG^0 of formation is $-39 \text{ kcal mol}^{-1}$ for yeast and $-38 \text{ kcal mol}^{-1}$ for mammals (at 37 °C in 1 M NaCl; see ref. 27). This suggested that the yeast and human U4/U6 RNA complexes would have similar melting temperatures. The T_m (temperature at which 50% dissociation is observed) of immunoaffinity-purified human U4/U6 snRNP, either as the intact particle or after treatment with SDS and proteinase K, has been determined to be $\sim 37^\circ\text{C}$ (ref. 2). We measured the T_m of the yeast *snR14/snR6* complex in proteinase-K/SDS-treated splicing extract Fraction I (ref. 28) under essentially identical ionic conditions (Fig. 4*a*). Surprisingly, the yeast complex was found to have a T_m of $\sim 53^\circ\text{C}$, 15 °C higher than human U4/U6. Approximately the same T_m was obtained for intact particles (Fig. 4*b*; the slightly lower T_m is probably due to the lower ionic strength of the buffer in the absence of SDS), indicating that the proteinase-K/SDS treatment does not affect the stability of the complex.

Although the T_m of a bimolecular RNA complex is concentration-dependent (because it reflects reassociation as well as dissociation, see ref. 27), this is very unlikely to explain the 15 °C

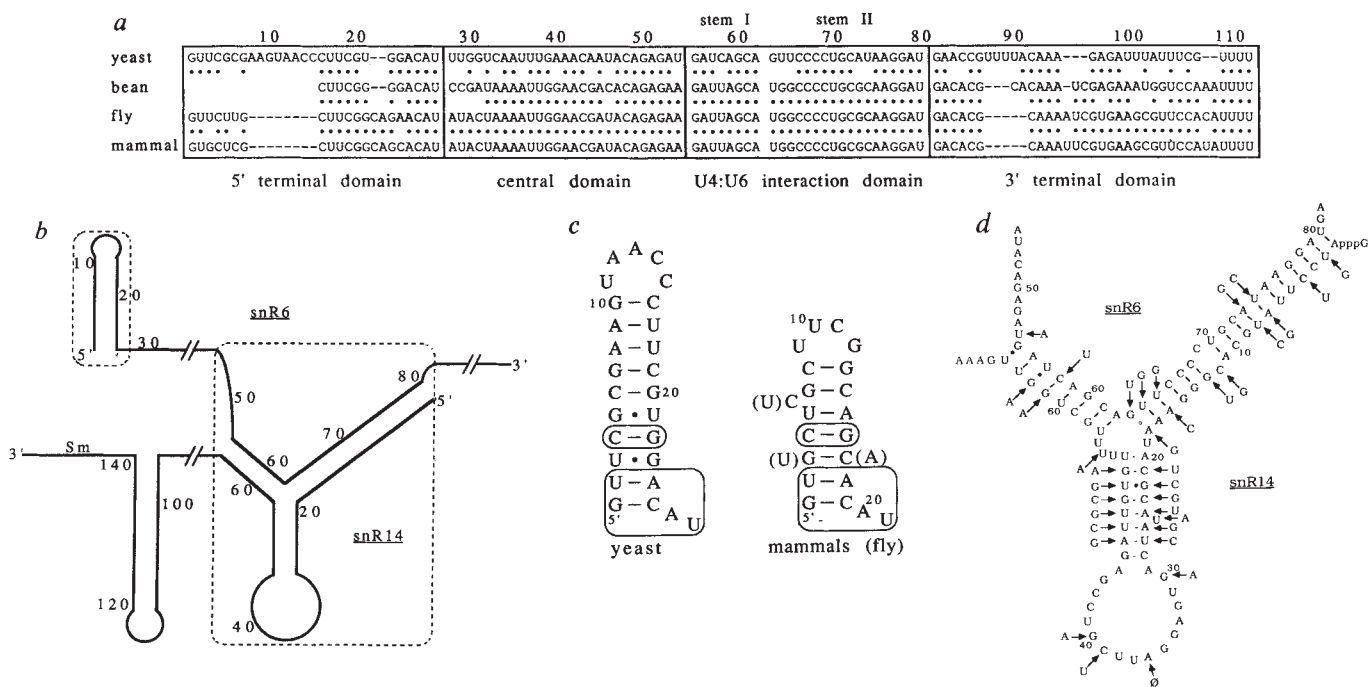


Fig. 3 *a*, Conservation of U6 primary structure. The nucleotide sequence of snR6 (yeast) is shown aligned with that of U6 RNA from *Vicia faba* L.⁴⁰ (bean), *Drosophila melanogaster*²⁰ (fruit fly), and mammal^{24,25,41}. The four domains (see text) are boxed, and sequences contained in U4:U6 stems I and II are labelled. Dashes represent gaps introduced to improve the alignment and dots indicate identical nucleotides. (We note that a small RNA from *Physarum* that was identified as U6 (ref. 42) initiates with a pyrimidine, cannot form the 5'-terminal stem, does not match the conservation pattern of other U6 RNAs and does not terminate in oligo(U). Therefore we consider it unlikely to be a U6 RNA.) *b*, Proposed structure of the yeast U4/U6 complex. snR6 is shown above snR14 (ref. 9). Adjacent parallel lines represent base paired stems. The regions elaborated in *c* and *d* are boxed. 'Sm' indicates the presumptive Sm antigen-binding site (ref. 9). *c*, Conservation of the 5'-terminal stem. The stem previously proposed²⁴ for mammalian U6 (with changes in fly in parentheses) is shown alongside its yeast homologue. Conserved nucleotides are boxed. *d*, Conservation of the U4/U6 interaction. Nucleotides 45-82 of yeast U6 (snR6) and 1-68 of yeast U4 (snR14) are shown arranged in the 'Y' structure described in the text. Substitutions found in mammalian U4 (refs. 43-45) and U6 are indicated with arrows.

difference. As the effect of concentration is less for RNAs with extensive intermolecular base-pairing, the yeast U4/U6 complex would have to be present at a very much higher concentration to generate the observed difference. This almost certainly was not the case, because the yeast extract was highly diluted before the determination, and because the concentration of snRNAs in the yeast nucleus appears to be much lower than that in the mammalian nucleus. To verify that the 15 °C difference was not due to some other aspect of our experimental protocol, we measured the T_m of the human U4/U6 complex in proteinase-K/SDS-treated HeLa cell nuclear extract by the same procedure. Unexpectedly, we observed both a low T_m and a high T_m form of the complex, with different mobilities on the native gel (Fig. 4c). The latter melted at about 55 °C, strikingly similar to the snR14/snR6 T_m . Neither complex hybridized to a probe complementary to U5 RNA (data not shown). The yield of the 55 °C complex varied with different extract preparations and occasionally was not detected at all, which might explain why it was not seen previously².

Our observation of a very stable form of the yeast and mammalian U4/U6 complex, as well as a less stable form of the mammalian complex², is particularly intriguing in light of recent studies that suggest the U4/U6 interaction is dynamic and intimately linked to spliceosome assembly. Native gel analysis of complexes formed when synthetic pre-mRNA is added to yeast^{10,29} or mammalian¹¹ *in vitro* splicing systems reveals that U4 RNA is lost during the transition from a pre-splicing complex to the functional spliceosome. U6 RNA, however, continues to comigrate with the spliceosome^{10,11}. These results suggest that the stability of the U4/U6 interaction is modulated during the process of splicing. It is possible that the high- T_m complex we

observed corresponds to the state of U4/U6 during the initial stages of assembly, and that the low- T_m species is an intermediate form of the fully destabilized U4/U6 complex of active spliceosomes. We are at present unable to explain the absence of a yeast low- T_m form in our experiments, or to account for the variability in the recovery of the HeLa cell high- T_m complex. Nevertheless, it seems likely that identification of the structural features that distinguish the different forms of the U4/U6 complex will lead to a better understanding of how their interaction might be modulated.

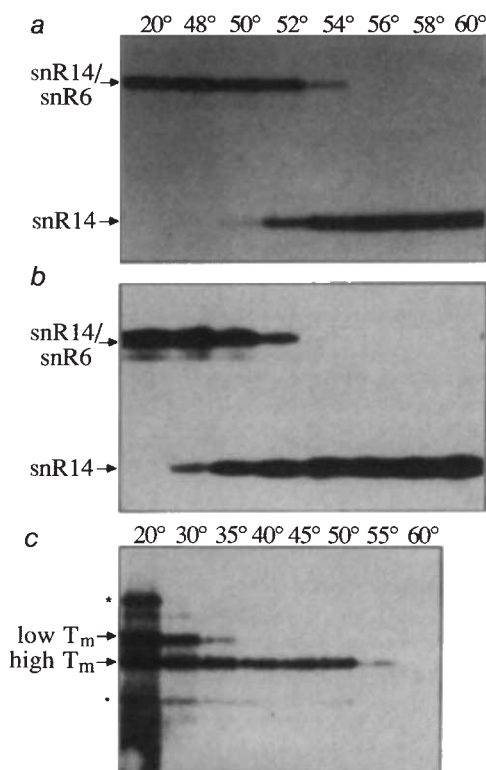
Conclusions

It has previously been shown that U6 RNA is important for splicing of pre-mRNA in mammalian cell extracts: oligodeoxynucleotide-directed RNase H cleavage of U6 results in a decreased efficiency of splicing³⁰. We have demonstrated that U6 is essential *in vivo*, and its high degree of sequence conservation implies a role fundamental to the splicing process. Moreover, the limited variation in the length of U6 between yeast and mammals is in sharp contrast to that exhibited by U1, U2 and U5 RNAs; conceivably this reflects a steric constraint imposed by a central position in the splicing complex. If this is the case, one would expect U6 to be in close contact with a number of other components of the assembled spliceosome.

U6 RNA can be divided into four domains (Fig. 3a), based on conserved structures. A stem/loop proposed for the 5'-terminal domain of mammalian U6 (refs 24, 25) can also be formed in yeast, with an unexpected feature: the sequence at the base of the stem is conserved, but its overall length and the sequence of the loop are variable (Fig. 3c). Interestingly, Black and Steitz³⁰ have shown that a dodecadeoxynucleotide com-

Fig. 4 *a*, T_m of the 'deproteinized' yeast U4/U6 complex. Aliquots of proteinase-K/SDS-treated Fraction I were incubated at the temperatures indicated in *a* for 5 min, then subjected to electrophoresis on a native gel. RNA was transferred to nylon membrane and probed for U4. The bands corresponding to free U4 (snR14) and its complex with U6 (snR14/snR6) are indicated. *b*, T_m of the yeast U4/U6 RNP. Aliquots of Fraction I were incubated at the temperatures indicated in *a* for 5 min, then quickly chilled. The samples were treated with proteinase K and SDS at 20 °C, then processed as in *a*. *c*, T_m of the human U4/U6 complex. Aliquots of proteinase-K/SDS-treated HeLa-cell nuclear-extract were treated as in *a*, and the membrane was probed for U6. Free U6 migrates with the buffer front and is not transferred to the membrane. A U4 probe hybridized to the high- and low- T_m bands, as well as to the band marked by an asterisk, which was not reproducibly seen. The minor, fast-migrating U6 form indicated with a dot contained no U4.

Methods. *a*, 2×10^8 cell equivalents (5 μ l) of yeast splicing extract Fraction I (ref. 28) was brought to 40 mM TrisCl, pH 7.4, 140 mM NaCl, 2% SDS, 2 mg ml⁻¹ proteinase K (E. Merck) in a total volume of 10 μ l and incubated at 20 °C for 1.5 h. This mixture was diluted 50-fold without changing the buffer composition (except for omission of proteinase K) and divided into eight 50- μ l aliquots, which were incubated as described above then frozen on dry ice. The samples were returned to 20 °C and immediately loaded on a 3%/8% polyacrylamide discontinuous native gel³³. Electrophoresis was at 7 V cm⁻¹ for 16 h at 4 °C. The gel was soaked for 1 h at 37 °C in 20 mM NaPO₄, pH 6.5, 8.3 M urea, 0.1% SDS, then for 1 h at 4 °C in 20 mM NaPO₄, pH 6.5. RNA was electroblotted from the separating gel to Hybond-N membrane (Amersham) in 20 mM NaPO₄, pH 6.5 at 7V for 21 h at 4 °C in a BioRad Trans-Blot cell. After UV treatment, the membrane was incubated with 6 $\times$ SSC (1 $\times$ SSC is 150 mM NaCl, 15 mM Na citrate), 5 $\times$ Denhardt's (1 $\times$ Denhardt's is 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin, 0.25 mM EDTA), 0.2% SDS for 2 h at 65 °C. Hybridization was in the same solution with ~ 100 ng (60×10^6 Cherenkov c.p.m.) each of 5'-[³²P]-labelled oligonucleotides 14B and 14C (ref. 9) at room temperature for 5.5 h. After washing for 30 min with three changes of 6 $\times$ SSC, 0.2% SDS at 37 °C, the membrane was exposed to Kodak XAR-5 film with a Cronex LP intensifying screen (Dupont) at -70 °C for 4 h. *b*, The temperature curve was carried out as in *a* except the 100-fold diluted Fraction I contained no SDS and was not treated with proteinase K. After quick-freezing, the 50- μ l aliquots were thawed at 20 °C and 5 μ l each of 20% SDS and 2 mg ml⁻¹ proteinase K were added. After 1 h at 20 °C, the samples were loaded on a native gel and processed as in *a*. *c*, 9×10^6 cell equivalents (10 μ l) of HeLa-cell nuclear-extract⁴⁶ (generously provided by M. Briggs and D. Black) was deproteinized as in *a* except the final volume was 100 μ l and the incubation was for 2.4 h. After an additional fivefold dilution, 50- μ l aliquots were incubated as described above. Electrophoresis, transfer and hybridization were as above but with 20×10^6 Cherenkov c.p.m. oligonucleotide 6A (5'-TC^A/TCTCTGTATTG-3') for 17 h. The membrane was washed as above and exposed for 17 h with a screen.



plementary to human U6, the 3' end of which pairs with the conserved ACAU at the base of this stem, can direct RNase H cleavage of U6 in splicing extract only when present at very high concentration, or after phenol extraction of the RNA. Thus the conserved region of the 5' terminal stem may constitute a recognition site for a U6 RNA-binding protein. Structures proposed for the central and 3' terminal domains^{24,25} cannot be formed in yeast. It is possible that U6 is using these apparently unstructured regions to interact with other factors that have not yet been identified. This seems especially likely for the central domain, which contains very highly conserved sequences.

The U4:U6 interaction domain of U6 RNA participates in two intermolecular stems. The yeast U4 (ref. 9) and U6 sequences provide the first phylogenetic verification of the previously proposed² stem II. The adjacent stem I, the existence of which was proven by chemical crosslinking of mammalian snRNAs²⁶, also exhibits compensatory changes in yeast. The high degree of sequence conservation in these stems suggests they are involved in additional interactions, either with proteins, or possibly with other RNAs when the U4/U6 complex is destabilized^{10,11}. In contrast, the U4 intramolecular stem between stems I and II is highly variable in sequence. The six compensatory base-pair changes in this stem between yeast and mammals provide incontrovertible proof of its existence. The recently sequenced trypanosome U4 and U6 RNAs (R. Nelson, personal communication) are significantly diverged from mammals and yeast, yet fit the trihelical 'Y' structure (Fig. 3d) very well. While *C. elegans* U4 and U6 RNAs are somewhat less diverged, it is notable that they can form a structure almost identical to the yeast and human Y structure (J. Thomas and T. Blumenthal, personal communication).

The contribution of individual base pairs to stability of the complex can now be tested by site-directed mutagenesis of *SNR14* (ref. 9) and *SNR6*. Mutants which exhibit cold-sensitive

(hyperstable) or heat-sensitive (hypostable) phenotypes would be expected. Such defects might be suppressed by mutations in genes coding for factors that modulate the U4/U6 interaction. The binding of such a factor (or its proteinase-K/SDS-resistant fragment) may be responsible for the slower migration of the human low- T_m complex in native gels. Alternatively, this could be the result of a large conformational alteration in the RNAs. Biochemical analysis of the complexes should allow us to discriminate between these two possibilities.

We thank Paul Siliciano for his contribution to the U4/U6 model, Tom Blumenthal, Rick Nelson, Mike Briggs and Doug Black for sharing data and materials, and Mary Ann Brow, Harry Noller, Ignacio Tinoco, Olke Uhlenbeck, Sandy Wolin, and members of our laboratory for discussions. This work was supported by grants from the NIH and NSF to C.G. D.A.B. is an American Cancer Society postdoctoral fellow.

Received 4 March; accepted 6 May 1988.

- Maniatis, T. & Reed, R. *Nature* **325**, 673-678 (1987).
- Hashimoto, C. & Steitz, J. A. *Nucleic Acids Res.* **12**, 3283-3293 (1984).
- Bringmann, P. et al. *EMBO J.* **3**, 1357-1363 (1984).
- Riedel, N., Wolin, S. & Guthrie, C. *Science* **235**, 328-331 (1987).
- Ares, M. *Cell* **47**, 49-59 (1986).
- Patterson, B. & Guthrie, C. *Cell* **49**, 613-624 (1987).
- Kretzner, L., Rymond, B. C. & Rosbash, M. *Cell* **50**, 593-602 (1987).
- Siliciano, P. G., Jones, M. H. & Guthrie, C. *Science* **237**, 1484-1487 (1987).
- Siliciano, P. G., Brow, D. A., Roiha, H. & Guthrie, C. *Cell* **50**, 585-592 (1987).
- Cheng, S.-C. & Abelson, J. *Genes Dev.* **1**, 1014-1027 (1987).
- Lamond, A. I., Konarska, M. M., Grabowski, P. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 411-415 (1988).
- Kunkel, G. R., Maser, R. L., Calvet, J. P. & Pederson, T. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8575-8579 (1986).
- Reddy, R., Henning, D., Das, G., Harless, M. & Wright, D. *J. biol. Chem.* **262**, 75-81 (1987).
- Das, G., Henning, D., Wright, D. & Reddy, R. *EMBO J.* **7**, 503-512 (1988).
- Carbon, P. et al. *Cell* **51**, 71-79 (1987).
- Kunkel, G. R. & Pederson, T. *Genes Dev.* **2**, 196-204 (1988).
- Bark, C., Weller, P., Zabielski, J., Janson, L. & Pettersson, U. *Nature* **328**, 356-359 (1987).
- Brow, D. A. *J. biol. Chem.* **262**, 13959-13965 (1987).
- Murphy, S., DiLiegro, C. & Melli, M. *Cell* **51**, 81-87 (1987).

20. Das, G., Henning, D. & Reddy, R. *J. biol. Chem.* **262**, 1187-1193 (1987).
21. Krol, A., Carbon, P., Ebel, J.-P. & Appel, B. *Nucleic Acids Res.* **15**, 2463-2478 (1987).
22. Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. *J. Proc. natn. Acad. Sci. U.S.A.* **78**, 6354-6358 (1981).
23. Boeke, J. D., Truehart, J., Natsoulis, G. & Fink, G. R. *Meth. Enzym.* **154**, 164-175 (1987).
24. Harada, F., Kato, N. & Nishimura, S. *Biochem. biophys. Res. Commun.* **95**, 1332-1340 (1980).
25. Reddy, R. & Busch, H. *Prog. Nucleic Acid Res. molec. Biol.* **30**, 127-162 (1983).
26. Rinke, J., Appel, B., Digweed, M. & Lührmann, R. *J. molec. Biol.* **185**, 721-731 (1985).
27. Freier, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9373-9377 (1986).
28. Cheng, S.-C. & Abelson, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2387-2391 (1986).
29. Pikielny, C. W., Rymond, B. C. & Rosbash, M. *Nature* **324**, 341-345 (1986).
30. Black, D. L. & Steitz, J. A. *Cell* **46**, 697-704 (1986).
31. Ohshima, Y., Okada, N., Tani, T., Itoh, Y. & Itoh, M. *Nucleic Acids Res.* **9**, 5145-5158 (1981).
32. England, T. E., Bruce, A. G. & Uhlenbeck, O. C. *Meth. Enzym.* **65**, 65-74 (1980).
33. Brow, D. A. & Geiduschek, E. P. *J. biol. Chem.* **262**, 13953-13958 (1987).
34. Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527-2538 (1977).
35. Lockhard, R. E. *et al. Nucleic Acids Res.* **5**, 37-56 (1978).
36. McPheeters, D. *et al. Nucleic Acids Res.* **14**, 5813-5826 (1986).
37. Broach, J. R. *Meth. Enzym.* **101**, 307-325 (1983).
38. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bact.* **153**, 163-168 (1983).
39. Rose, M. D. *Meth. Enzym.* **152**, 481-504 (1987).
40. Kiss, T., Antal, M. & Solymosy, F. *Nucleic Acids Res.* **15**, 543-560 (1987).
41. Sri-Widada *et al. Molec. Biol. Rep.* **8**, 29-36 (1981).
42. Skinner, H. B. & Adams, D. S. *Nucleic Acids Res.* **15**, 371 (1987).
43. Reddy, R., Henning, D. & Busch, H. *J. biol. Chem.* **256**, 3532-3538 (1981).
44. Kato, N. & Harada, F. *Biochem. biophys. Res. Commun.* **99**, 1477-1485 (1981).
45. Krol, A. *et al. Nucleic Acids Res.* **9**, 2699-2716 (1981).
46. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).

A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs

Claudio Santoro^{*†}, Nicolas Mermoud^{*}, Philip C. Andrews[†] & Robert Tjian^{*}

^{*} Howard Hughes Medical Institute, Department of Biochemistry, University of California, Berkeley, California, USA

[†] Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907, USA

The CTF/NF-I group of cellular DNA binding proteins recognizes the sequence GCCAAT and is implicated in eukaryotic transcription as well as DNA replication. Molecular analysis of human CTF/NF-I complementary DNA clones reveals multiple messenger RNA species containing alternative coding regions, apparently as a result of differential splicing. Expression and functional analysis establish that individual gene products can bind to GCCAAT recognition sites and serve both as promoter-selective transcriptional activators and as initiation factors for DNA replication.

IN eukaryotic cells, regulatory regions of DNA responsible for directing transcription and replication contain multiple recognition elements that interact with sequence-specific DNA-binding proteins. Until recently, it was not known whether cellular *trans*-acting factors that regulate transcription in eukaryotic cells were related to proteins involved in initiating DNA replication. However, analysis of protein-DNA interactions at eukaryotic promoters and origins of DNA replication has suggested that in several instances the same cellular site-specific DNA-binding proteins can participate in both transcription and replication events^{1,2}. One of the first documented cases involved a factor isolated from human cells that binds the DNA sequence GCCAAT².

The GCCAAT recognition motif, often referred to as the CCAAT-box, was originally identified as an important RNA polymerase II promoter element in the vertebrate globin genes and in a variety of other eukaryotic genes³⁻¹³. Purification of human GCCAAT-binding factors by sequence-specific DNA-affinity chromatography identified a series of polypeptides ranging in relative molecular mass (M_r) from 52,000 (52K) to 66,000 (66K) (ref. 2). *In vitro* transcription studies using various templates, including the HSV-TK and α -globin promoters established that these DNA-binding proteins, collectively designated CCAAT-binding transcription factors (CTF), act as *bona fide* sequence-specific transcription initiation proteins^{2,14}.

In addition to being associated with various RNA polymerase II promoters, the GCCAAT element can also serve as a *cis* control sequence in viral origins of DNA replication. A group of cellular GCCAAT-binding proteins originally designated nuclear factor I (NF-I) directs initiation of Adenovirus DNA synthesis by binding to specific origin sequences¹⁵⁻¹⁸. Subsequent purification and biochemical characterization established that the CTF and NF-I group of polypeptides are indis-

tinguishable in their DNA-binding properties, immunological cross-reactivity, polypeptide composition, proteolytic cleavage products and *in vitro* activation of transcription and DNA replication (refs 2, 19; C.S., unpublished data). These studies suggested that the CTF/NF-I polypeptides represent a class of factors that have evolved to initiate different DNA transactions such as transcription and replication, perhaps by related mechanisms involving the binding of CCAAT motifs. But it is not known whether individual CCAAT binding polypeptides activate both transcription and initiation of DNA replication or whether functionally distinct members of the CTF/NF-I proteins carry out each of the two activities. The multiple CTF/NF-I proteins could be the products of several related genes each encoding a functionally different version of the CCAAT-binding protein, or the different proteins could be products of a single gene that have undergone different post-translational modifications, giving rise to distinct polypeptides with altered functions. A third possibility is that a single gene encodes multiple CTF/NF-I products that are generated by alternative splicing of mRNAs.

Here we report the isolation of multiple human cDNA species encoding different members of the CTF/NF-I family of proteins. Our findings indicate that different CTF/NF-I polypeptide species probably arise as a consequence of differential RNA splicing. Gene products encoded by three of these human cDNA species were expressed in *Escherichia coli* and all were found to be CCAAT-binding proteins. Also we have established that single CTF/NF-I products can serve both as a promoter selective transcription factor for RNA polymerase II and as an initiation factor for adenovirus DNA replication.

Results

Isolation of multiple CTF cDNAs containing alternative coding regions. HeLa cell CTF polypeptides were purified from crude nuclear extracts by gel filtration and sequence-specific DNA-affinity chromatography². The amino-acid sequence of five

[†] Present address: Dipartimento di Genetica, Biologia e Chimica Medica, Università di Torino, 10126 Torino, Italy

purified tryptic-digest peptides was subsequently determined by sequential Edman degradation. A 33-nucleotide synthetic oligomer corresponding to the expected coding sequence of CTF peptide 1 (PEP1; KELDLYLAYFVR) was selected as a specific DNA probe for screening a HeLa cDNA library, and six independent cDNA clones were isolated from 10^6 λ gt10 phage plaques. The DNA sequence of the longest cDNA isolate contained 1.75 kilobases (kb) of human DNA and was designated CTF-1. The entire sequence of CTF-1 was determined by the M13 dideoxy sequencing technique (Fig. 1). Inspection of the CTF-1 cDNA sequence shows a long, open reading frame encompassing the first 1,566 bases. The amino-acid sequences of all five tryptic peptides are preceded either by an arginine or lysine residue as expected for fragments generated by trypsin digestion. Direct alignment of the deduced amino-acid sequence of CTF-1 with the sequence of the five tryptic peptide fragments (PEP1 through PEP5) of purified HeLa CTF proteins confirmed the correct reading frame and the authenticity of the cDNA clone (Fig. 1).

The remaining five cDNA isolates obtained from the original screen were analysed along with four new isolates recovered from a secondary screen of 0.5×10^6 phage plaques using the CTF-1 cDNA clone as a probe. Restriction endonuclease cleavage patterns (data not shown) and partial DNA sequence of these nine independently isolated cDNA clones revealed three structurally related but distinct classes of cDNA isolates. A direct comparison of these three types of cDNAs shows that each consists of identical blocks of coding regions interspersed with two stretches that contain alternative coding regions (Fig. 1a, b). These findings may, in part, explain the presence of multiple polypeptide forms of CTF purified from HeLa cells and suggests that alternative splicing generates a family of CTF proteins.

A restriction fragment common to all three CTF cDNA isolates was used as a probe in Southern blot analysis and under stringent hybridization conditions, only one human genomic DNA fragment was detected after digestion with several restriction enzymes (Fig. 2a). Thus, it appears that the different CTF cDNAs isolated thus far are encoded by a single gene in human cells. Hybridization of HeLa cell poly-A-selected cytoplasmic RNA with a probe containing the entire CTF-1 cDNA identified two mRNA forms consisting of a major species of 8.6 kb and a minor species of 4.5 kb (Fig. 2b). However, primer extension analysis using a synthetic probe derived from the 5' end of CTF cDNA sequences (nucleotides 88–120) (see Fig. 1b) gave rise to a unique extension product of 240 ± 2 nucleotides (Fig. 2c). Thus, the multiple CTF mRNA species contain various 3'-terminal portions and a common 5' end that extends 120 nucleotides beyond the longest partial cDNA sequence obtained thus far. These results, taken together with the deduced amino-acid sequence of CTF-1, -2 and -3, and the presence of a consensus translational start signal (see Fig. 1b) suggests that the methionine triplet at nucleotide 68/70 is likely to represent the initiation codon. However, as the open reading frame continues to the 5' end of the CTF-1 sequence, we cannot exclude that the coding sequence may extend past the proposed initiation codon. Isolation of the genomic region encoding CTF and detailed structural analysis will be necessary to confirm this observation.

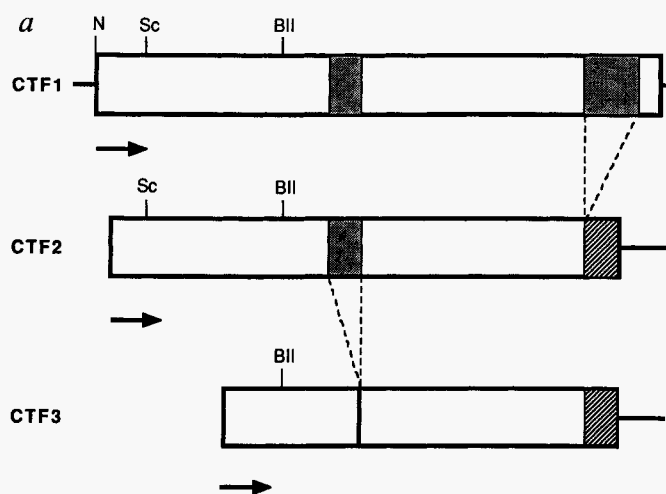
Expression and DNA-binding properties of CTF proteins. The entire coding region of CTF-1 was inserted into a bacterial vector (pBS) to allow expression of a hybrid protein containing 499 amino acids of CTF fused to 34 amino acids of vector sequences at the N terminus (Fig. 3a). This bacterial expression vector is designated pCTF-1. We also constructed two additional expression vectors (pCTF-2 and -3) by substituting the coding regions of pCTF-1 with sequences from CTF-2 and CTF-3 (Fig. 3a). Extracts from bacteria transformed with pCTF-1, -2 and -3 or control pBS vector DNA were prepared and subjected to Western blot analysis, using antibodies directed against CTF/NF-I polypeptides² (Fig. 3b). Proteins of ~62, 58 and 55K

corresponding to the largest fusion products of pCTF-1, -2 and -3, respectively, were found to cross-react specifically with anti-CTF/NF-I antibodies. Also, we observed several smaller species of CTF-1, -2 and -3 crossreacting with anti-NF-I sera that are probably generated by proteolytic degradation or initiation of translation from internal sites.

To investigate the DNA-binding specificity and other biochemical properties of CTF expressed in bacteria, we have purified the three CTF fusion proteins by a combination of conventional and DNA-affinity chromatography. 5' end-labelled DNA fragments containing CTF/NF-I binding sites derived from α -globin²⁰, Adenovirus origin of replication¹⁶, and h-ras sequences²¹ were used as templates for DNase I protection assays². A comparison of the DNA-binding specificities of purified HeLa CTF polypeptides with those of CTF fusion proteins encoded by pCTF-1, -2 and -3 shows that the DNase I footprint patterns obtained are essentially identical in all cases (Fig. 4). A few minor alterations in the enhanced cleavage pattern can be seen which likely reflect subtle differences between the HeLa proteins and bacterial fusion polypeptides. The bacterial CTF proteins also binds the relatively low affinity CCAAT-box element in the HSV-TK promoter with an efficiency indistinguishable from that of HeLa CTF/NF-I (data not shown). These results provide direct evidence that each of the three different classes of cDNAs encode a GCCAAT-binding protein. Moreover, our findings indicate that a single CTF polypeptide can bind its cognate recognition sequence without the aid of related CTF proteins and that the DNA binding specificity of several individual members of the CTF/NF-I family is indistinguishable from the collective binding specificity of the entire family of polypeptides.

Activation of transcription by CTF-1 and -2. One important issue that has remained unresolved concerning the biochemical function of the CTF family of polypeptides is whether a single polypeptide species can carry out transcriptional activation of promoters containing GCCAAT elements. A convenient and powerful experimental approach would be to develop a complementation assay in which a cell line lacking functional CTF could be transfected with wild-type or mutant copies of the CTF gene. At present, there is no clear indication that a mammalian cell line exists which does not express at least some CTF/NF-I protein and there are no known cell lines carrying conditional lethal CTF genes. We have recently discovered that the general transcriptional apparatus of *Drosophila* cells is compatible with the action of mammalian transcription factors and that *Drosophila* cells lack promoter selective factors such as Sp1 and CTF, which are ubiquitous in vertebrate cells (A. Courey and R.T., manuscript in preparation; C.S., unpublished results). In the absence of an appropriate mammalian host, *Drosophila* cells such as the Schneider line has provided a useful *in vivo* assay system for studying mammalian transcription factors (A. Courey and R.T., submitted) and we have adopted such an approach for determining the functional specificity of the individual CTF products.

First, we used plasmid vectors (pPADH) to direct expression of CTF proteins from the tandem promoters of the *Drosophila* alcohol dehydrogenase gene²². The coding regions of CTF-1 and -2 were inserted into the pPADH vector (Fig. 5a) to produce CTF fusion products with the first methionine of the *Adh* gene at the N terminus. These plasmid vectors were designated pPADH-CTF-1 and pPADH-CTF-2. Next, we prepared reporter plasmids containing the bacterial chloramphenicol acetyltransferase gene (CAT) linked to the human α -globin promoter (p α -CAT- Δ 87) or a truncated version (p α -CAT- Δ 55) lacking the GCCAAT element. Transfection of either p α -CAT- Δ 87 or Δ 55 into Schneider cells alone or with the pPADH control plasmid results in a low basal level of CAT activity (Fig. 5b). By contrast, when pPADH-CTF-1 or -2 is co-transfected with p α -CAT- Δ 87, there is a significant induction (3–5-fold) of CAT activity. But, expression of either CTF-1 or -2 did not induce



b

CTF-1 GAATTCGCG GCGGCGCG AGCGGCTCG GTCCCGCG CCGGCTCG CTCCTCGAG CAGCGC 67

CTF-2

CTF-3

CTF-1 ATG GAT GAG TTC CAC CCG TTC ATC GAG GCC CTG CTG CCT CAC GTC CCG GCC TTC GCC 124

CTF-2 MET Asp Glu Phe His Pro Phe Ile Glu Ala Leu Leu Pro His Val Arg Ala Phe Ala

CTF-3

CTF-1 TAC ACC TGG TTC AAC CTG CAG GCG CCG AAG CGC AAG TAC TTC AAG AAG CAC GAG AAG 181

CTF-2 Tyr Thr Trp Phe Asn Leu Gln Ala Arg Lys Arg Lys Tyr Phe Lys Lys His Glu Lys

CTF-3

CTF-1 CCG ATG TCG AAG GAC GAG GAG CGT GCG GTC AAG GAC GAG CTG CTG GCG GAG AAG CCC 238

CTF-2 Arg MET Ser Lys Asp Glu Glu Arg Ala Val Lys Asp Glu Leu Leu Gly Glu Lys Pro

CTF-3

CTF-1 GAG GTC AAG CAG AAG TGG GCG TCG CCG CTG CTG GCC AAG CTG CCG AAG GAC ATC CCG 295

CTF-2 Glu Val Lys Gln Lys Trp Ala Ser Arg Leu Leu Ala Lys Leu Arg Lys Asp Ile Arg

CTF-3

CTF-1 CCC GAG TGC CCG GAG GAC TTC GTG CTG AGC ATC ACC GGC AAG AAG GCG CCG GCG TGC 352

CTF-2 Pro Glu Cys Arg Glu Asp Phe Val Leu Ser Ile Thr Gly Lys Lys Ala Pro Gly Cys

CTF-3

CTF-1 GTG CTC TCC AAC CCC GAC CAG AAG GGC AAG ATG CCG CCG ATC GAC TGT CTC CCG CAG 409

CTF-2 Val Leu Ser Asn Pro Asp Gln Lys Gly Lys MET Arg Arg Ile Asp Cys Leu Arg Gln

CTF-3

CTF-1 GCG GAC AAG GTG TGG CCG CTG GAC CTG GTC ATG GTC ATC CTG TTC AAG GGC ATC CCG 466

CTF-2 Ala Asp Lys Val Trp Arg Leu Asp Leu Val MET Val Ile Leu Phe Lys Gly Ile Pro

CTF-3

CTF-1 CTG GAG AGC ACC GAC GGC GAG CCG CTG GTC AAG GCT GCG CAG TGC GGT CAC CCG GTC 523

CTF-2 Leu Glu Ser Thr Asp Gly Glu Arg Leu Val Lys Ala Ala Gln Cys Gly His Pro Val

CTF-3

CTF-1 CTG TGC GTG CAG CCG CAC CAC ATT GGC GTG GCC GTC AAG GAG CTG GAC CTC TAC CTG 580

CTF-2 Leu Cys Val Gln Pro His His Ile Gly Val Ala Val Lys Glu Leu Asp Leu Tyr Leu

CTF-3

CTF-1 GCG TAC TTC GTG CGT GAG CGA GAT GCA GAG CAA AGC GGC AGT CCC CCG ACA GCG ATG 637

CTF-2 Ala Tyr Phe Val Arg Glu Arg Asp Ala Glu Gln Ser Gly Ser Pro Arg Thr Gly MET

CTF-3

CTF-1 GGC TCT GAC CAG GAG GAC AGC AAG CCC ATC ACG CTG GAC ACG ACC GAC TTC CAG GAG 694

CTF-2 Gly Ser Asp Gln Glu Asp Ser Lys Pro Ile Thr Leu Asp Thr Thr Asp Phe Gln Glu

CTF-3

CTF-1 AGC TTT GTC ACC TCC GGC GTG TTC AGC GTC ACT GAG CTC ATC CAA GTG TCC CCG ACA 751

CTF-2 Ser Phe Val Thr Ser Gly Val Phe Ser Val Thr Glu Leu Ile Gln Val Ser Arg Thr

CTF-3

CTF-1 CCC GTG GTG ACT GGA ACA GGA CCC AAC TTC TCC CTG GGG GAG CTG CAG GGG CAC CTG 808

CTF-2 Pro Val Val Thr Gly Thr Gly Pro Asn Phe Ser Leu Gly Glu Leu Gln Gly His Leu

CTF-3

CTF-1 GCA TAC GAC CTG AAC CCA GCC AGC ACT GGC CTC AGA AGA ACG CTG CCC AGC ACC TCC 865

CTF-2 Ala Tyr Asp Leu Asn Pro Ala Ser Thr Gly Leu Arg Arg Thr Leu Pro Ser Thr Ser

CTF-3

CTF-1 TCC AGT GGG AGC AAG CCG CAC AAA TCG GGC TCG ATG GAG GAA GAC GTG GAC ACG AGC 922

CTF-2 Ser Ser Gly Ser Lys Arg His Lys Ser Gly Ser MET Glu Glu Asp Val Asp Thr Ser

CTF-3

CTF-1 CCT GGC GGC GAT TAC TAC ACT TCG CCC AGC TCG CCC ACG AGT AGC AGC CCG AAC TGG 979

CTF-2 Pro Gly Gly Asp Tyr Tyr Thr Ser Pro Ser Ser Pro Thr Ser Ser Ser Arg Asn Trp

CTF-3

CTF-1 ACG GAG GAC ATG GAA GGA GGC ATC TCG TCC CCG GTG AAG AAG ACA GAG ATG GAC AAG 1036

CTF-2 Thr Glu Asp MET Glu Gly Gly Ile Ser Ser Pro Val Lys Lys Thr Glu MET Asp Lys

CTF-3

CTF-1 TCA CCA TTC AAC AGC CCG TCC CCC CAG GAC TCT CCC CCG CTC TCC AGC TTC ACC CAG 1093

CTF-2 Ser Pro Phe Asn Ser Pro Ser Pro Gln Asp Ser Pro Arg Leu Ser Ser Phe Thr Gln

CTF-3

CTF-1 CAC CAC CCG CCC GTC ATC GCC GTG CAC AGC GGG ATC GCG CCG AGC CCA CAC CCG TCC 1150

CTF-2 His His Arg Pro Val Ile Ala Val His Ser Gly Ile Ala Arg Ser Pro His Pro Ser

CTF-3

CTF-1 TCC GCT CTG CAT TTC CCT AGC ACG TCC ATC CTA CCC CAG ACG GCC TCC ACC TAC TTC 1207

CTF-2 Ser Ala Leu His Phe Pro Thr Thr Ser Ile Leu Pro Gln Thr Ala Ser Thr Tyr Phe

CTF-3

CTF-1 CCC CAC ACG GCC ATC CCG TAC CCA CCT CAT CTC AAC CCC CAG GAC CCG CTC AAA GAT 1264

CTF-2 Pro His Thr Ala Ile Arg Tyr Pro Pro His Leu Asn Pro Gln Asp Pro Leu Lys Ser

CTF-3

CTF-1 CTT GTC TCG CTG GCC TGC GAC CCA GCC AGC CAG CAA CCT GGA CCG TTA AAT GGA AGT 1321

CTF-2 Leu Val Ser Leu Ala Cys Asp Pro Ala Ser Gln Gln Pro Gly Pro Leu Asn Gly Ser

CTF-3

CTF-1 GGT CAG CTC AAA ATG CCC AGC CAC TGC CTT TCT GCT CAG ATG CTG GCA CCT CCG CCC 1378

CTF-2 Gly Gln Leu Lys MET Pro Ser His Cys Leu Ser Ala Gln MET Leu Ala Pro Pro Pro

CTF-3

CTF-1 CCG GGG CTG CCA CCG CTG GCG CTC CCC CCT GCG ACC AAA CCC GCG ACC ACC TCC GAG 1435

CTF-2 Pro Gly Leu Pro Arg Leu Ala Leu Pro Pro Ala Thr Lys Pro Ala Thr Thr Ser Glu

CTF-3

CTF-1 GGA GGA GGC ACG TCG CCG ACC TCG CCT TCC TAT TCT CCG CCC GAC ACG TCC CCT GCA 1492

CTF-2 Gly Gly Ala Thr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Pro Asp Thr Ser Pro Ala

CTF-3 Pro Thr Leu Arg Pro Thr Arg Pro Leu Gln

CTF-1 AAC CGT TCC TTT GTG GGA TTA GGA CCA AGG GAT CCT GCG GGC ATT TAT CAG GCA CAG 1549

CTF-2 Asn Arg Ser Phe Val Gly Leu Gly Pro Arg Asp Pro Ala Gly Ile Tyr Gln Ala Gln

CTF-3 Thr Val Pro Leu Trp Asp

CTF-1 TCC TGG TAT CTG GGA TAG CAAAGGTCTT CTTCCTCGC CCCTTCTCA TCGTCCAGG AATCCAGGG 161

CTF-2 Ser Trp Tyr Leu Gly

CTF-3

CTF-1 GGCAGCACAG CCGCCCCCG CCCACGTTTT TGSTGGAAAA TTAGAGTGAA CAAGAACCCT CCGCCACT 1687

CTF-2

CTF-3

CTF-1 CCCAGCCCG CCAAAAAGAC AAAACACATA GACGCACACA CTCAGGAGGA AAGAAAAAC CGAATTC 1755

CTF-2

CTF-3

expression of CAT activity from α -CAT- Δ 55, which lacks a CTF binding site. These findings indicate that both CTF-1 and -2 encode transcription factors that act in a sequence-dependent manner. Our results also establish that individual CTF proteins can activate α -globin transcription from a template containing a GCCAAT motif in *Drosophila* cells, even in the absence of other mammalian specific promoter or enhancer factors. But the induction of α -globin transcription might be more dramatic in the presence of other mammalian factors that interact, perhaps synergistically, with CTF proteins to potentiate transcription.

Activation of Adenovirus DNA replication by CTF proteins. Previous studies suggested that the collection of proteins designated CTF/NF-I are responsible for activating both transcription and DNA replication. But it was not possible to demonstrate that any single protein was capable of carrying out both the transcription and replication functions. The isolation of cDNA clones encoding three different classes of CTF products now allows us to address this issue in a direct way. An early step in the replication of Adenovirus involves the covalent attachment of the viral coded precursor terminal protein (pTP) to the first

Fig. 1 DNA and amino-acid sequence of HeLa CTF. *a*, Schematic representation of three partial cDNA clones encoding HeLa CTF. The coding regions common to CTF-1, -2, -3 are shown as open rectangles. The stippled boxes represent alternative coding regions present in CTF-1 but absent in CTF-2 and/or CTF-3. Hatched boxes designate an open reading frame encoding amino acids at the C-terminal end of CTF-2 and -3 that are distinct from CTF-1 sequences and are the result of a frameshift at the 3' end of the cDNA sequences. The direction of transcription is indicated by thick horizontal arrows. Cleavage sites in CTF-1, -2, -3 cDNAs for restriction enzymes *Nco*I (N), *Sca*I (Sc) and *Bst*EII (BII) are demarcated by vertical bars. *b*, The entire nucleotide sequence of CTF-1 cDNA is shown along with the deduced amino-acid sequence. The amino-acid sequences enclosed by boxes (designated PEPI-5) represent the five tryptic peptides derived from HeLa CTF. The sequences of CTF-2 and -3 that are identical to CTF-1 are represented by thick stippled and hatched lines, respectively. The alternative coding regions absent from CTF-2 and/or -3 are shown as thin dotted lines. The predicted C-terminal amino-acid sequence of CTF-2 and -3 that differs from that of CTF-1 is shown in italics. Stop codons are shown by asterisks.

Methods. CTF protein was purified from HeLa extracts by sequence-specific DNA-affinity chromatography as described previously². About 100 µg (1.5 nmol) of intact CTF was subjected to cyanogen bromide cleavage at 100-fold molar excess CNBr over CTF in 88% formic acid in the dark at room temperature for 18 hours. Trypsin digestion of cyanogen bromide-cleaved CTF was performed by incubation in 1.0 ml of 1.1 M urea, 0.08 M *N*-ethylmorpholinium acetate, pH 8.0 at 37 °C for 3 h with 2.5 µg affinity-purified trypsin. A second addition of 2.5 µg trypsin was then made and the digestion continued for 3 more hours. The digest was then reduced by addition of 100 µl of 10% tributylphosphine⁴¹ followed by incubation under nitrogen at 37 °C for 45 min. The reduced digest was acidified by addition of 50 µl trifluoroacetic acid and then applied to a VYDAC C8 (the separations group) pH-stable reversed-phase HPLC (high pressure liquid chromatography) column and the peptides eluted at 0.5 ml min⁻¹ by a gradient from 0 to 50% acetonitrile in 60 min. Both solvents contained 0.1% trifluoroacetic acid. The elution profile of the HPLC column was monitored by absorbance at 215 nm. A parallel control digest was also run to identify artefact peaks arising from trypsin fragments. Automated Edman degradation of the tryptic peptides was performed on an Applied Biosystems model 470A with off-line analysis of PTH (phenylthiohydantoin) amino acids as described elsewhere⁴². An oligonucleotide, 5'-ACGCACGAAGTAGGCIAGGT-AIAGGTCIAGCTC was deduced from the sequence of peptide 1 (see Fig. 1) according to Lathe⁴³ (I represents inosine) and used to screen a HeLa cDNA library in λgt10 kindly provided by Dr C. Hauser. Lambda plaques transferred to nitrocellulose were hybridized in a mixture containing 6×SSC (1×SSC is 150-mM NaCl, 15-mM sodium citrate), 50-mM Na-phosphate pH 7, 5× Denhardt's 100 µg ml⁻¹ sonicated salmon sperm DNA and 1 pmol ml⁻¹ of ³²P-labelled oligonucleotide. Hybridization of the filters was carried out at 37 °C for 24 h, washed once at room temperature and twice at 50 °C for 15 min each in 2×SSC, 0.5% sodium dodecyl sulphate (SDS) and subjected to autoradiography at -70 °C with intensifying screens. Six independent clones of CTF-1 to -6 were isolated from 10⁶ plaques. CTF-1 cDNA was used as probe to screen a human placenta cDNA library under stringent hybridization conditions and four additional independent clones were isolated, CTF-7 to -10, from 5×10⁵ plaques. All CTF cDNA clones were analysed by restriction mapping and dideoxy sequencing⁴⁴. CTF-1 and -3 were completely sequenced, and CTF-2 was partially sequenced in the regions which differ from CTF-1.

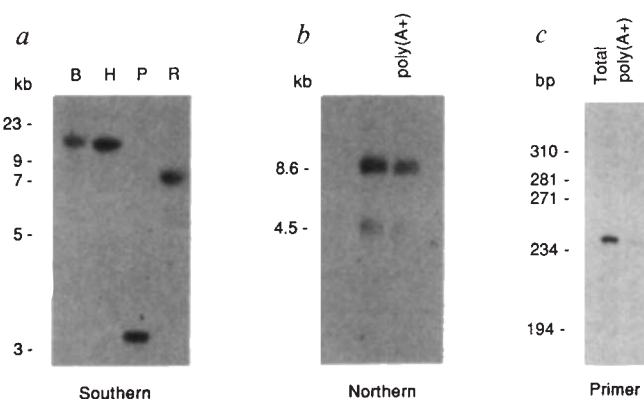


Fig. 2 Messenger RNA and gene structure of CTF. *a*, Southern blot analysis of human genomic DNA. Human placenta DNA (10 µg) was digested with *Bam*HI (B), *Hind*III (H), *Pst*I (P) and *Eco*RI (R) and fractionated by agarose gel (1%) electrophoresis. After transfer to gene screen membrane, the DNA was hybridized to a *Bam*HI-*Eco*RI fragment encompassing the C-terminal coding sequence and the first portion of the 3'-untranslated region of CTF-1 containing nucleotides 1,530 to 1,755 (see Fig. 1). Hybridization was carried out in 15% formamide, 0.2 M sodium phosphate pH 6.5, 7% SDS, 1 mM EDTA, 1 g 100 ml⁻¹ bovine serum albumin with 10⁷ counts min⁻¹ of ³²P-labelled probe at 65 °C for 15 h. The filters were washed twice at 68 °C in 0.1×SSC, 0.5% SDS. Relative molecular mass markers consisted of the 1-kb ladder from Bethesda Research Laboratories. *b*, Northern blot analysis of CTF mRNA isolated from HeLa cells. Cytoplasmic RNA was extracted from HeLa cells after lysis with NP40, and poly(A)⁺ RNA was enriched by standard procedures. Northern blots were prepared as described previously⁴⁵ and hybridized to a probe consisting of the entire cDNA (*Eco*RI fragment) of CTF-1, ³²P-labelled by random primed synthesis. Hybridization was carried out as described previously⁴⁶ and filters were washed twice in 68 °C for 15 min each in 0.2×SSC, 0.5% SDS. *c*, Localization of the 5' end of CTF mRNA by primer extension analysis. About 50 µg of total RNA or 5 µg of poly(A)⁺ selected RNA from HeLa cells were subjected to primer extension reactions as described previously⁴⁵. A synthetic oligonucleotide complementary to nucleotides 88 to 120 (see Fig. 1) was used as primer. Synthesis of DNA was carried out at 42 °C for 1 h as described previously⁴⁶.

HeLa cell CTF/NF-I as well as the products of pCTF-1 or -2 purified from bacteria. Control proteins derived from bacteria lacking pCTF DNA showed no stimulation of pTP/dCMP complex formation. These data indicate that the products of CTF-1 and -2 are capable of serving as initiation factors for DNA replication and clearly establishes that individual CTF proteins can carry out both transcriptional activation and DNA replication functions.

Discussion

Structure and function of CCAAT-box-binding factors. We previously identified a group of regulatory proteins in human cells (CTF/NF-I) that recognizes GCCAAT elements and activates both transcription and DNA replication in cell-free reactions². Here, we describe the molecular cloning of multiple cDNA isolates encoding members of the CTF/NF-I family of human CCAAT-box-binding proteins. Structural analysis of ten independent cDNA clones suggests that a single gene in human cells gives rise to at least three distinct mRNA species which are most likely generated by alternative RNA splicing. It seems likely, therefore, that at least some of the distinct forms of CTF proteins purified from HeLa cells are encoded by differentially processed RNA transcripts. We cannot, however, eliminate the possibility that the mixture of CTF polypeptides in HeLa cells also includes post-translationally modified versions of these three forms of CTF protein.

Several lines of evidence establish that, although the three

deoxynucleotide residue (a cytidine-5' monophosphate) of the newly synthesized DNA strand. The formation of this pTP-dCMP initiation complex in cell-free reactions requires the presence of the Adenovirus DNA polymerase, viral origin sequences and several nuclear proteins that interact with the origin DNA, including CTF/NF-I polypeptides²³. Reconstitution reactions using cellular fractions (BR-FT) containing various initiation factors but depleted of CTF/NF-I allow a low level of pTP/dCMP complex to form. However, the initiation reaction is greatly stimulated (10×) by the addition of purified

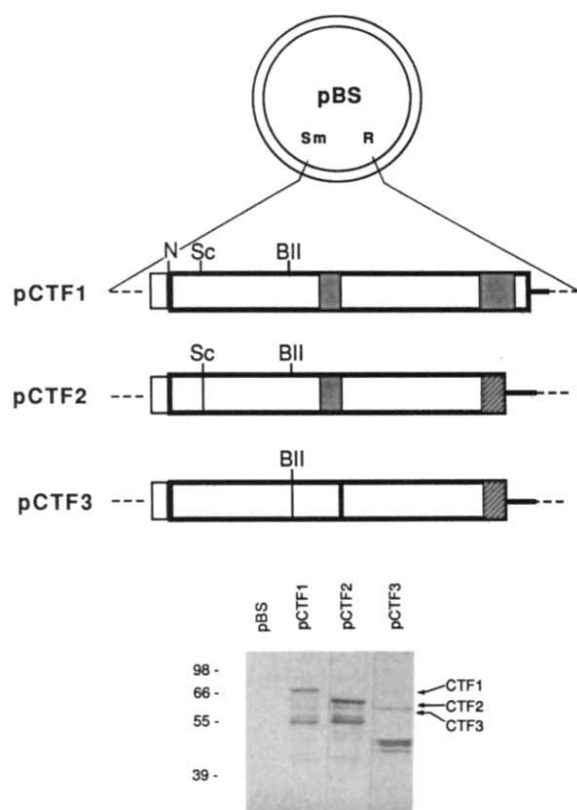


Fig. 3 Expression of CTF proteins in bacteria. *a*, Schematic diagram of vectors used for expression of CTF cDNAs in bacteria. The coding sequences are shown as open rectangles. The alternative coding regions are designated by stippled or hatched boxes. The first 34 codons of the β -galactosidase of the pBSSK⁺ vector (Stratagene) that are fused to the N terminus of the CTF cDNAs are represented by small rectangles at the 5' ends of the cDNA. These expression vectors were used to transform *E. coli* JK50 (ref. 47) and synthesis of the fusion proteins encoded by pCTF-1, -2, and -3 were induced by addition of IPTG (isopropyl- β -D-thiogalactopyranoside). *b*, Western blot analysis of crude bacterial cell lysates from JK50 cells transformed with either vector sequences alone (pBS) or vector sequences carrying pCTF after induction with IPTG. The full length CTF products (indicated by arrows) and various degradation products were detected by immunoblot analysis using rabbit anti-NF serum². Relative-molecular-mass markers are indicated on the left of panel *b*.

Methods. The pCTF-1 expression vector was constructed by digesting CTF-1 cDNA with endonuclease *Nco*I and *Eco*RI. The *Nco*I site was repaired with Klenow polymerase and the resulting restriction enzyme fragment was inserted into vector pBSSK⁺, cleaved with *Sma*I (Sm) and *Eco*RI (R). pCTF-2 was obtained by the insertion of a 1.4 kb *Sca*I-*Eco*RI fragment from CTF-2 (see Fig. 1a) in pCTF-1 after digestion with *Sca*I and *Eco*RI. pCTF-3 was obtained by inserting a 1.1-kb *Bst*EII-*Eco*RI fragment from CTF-3 into pCTF-1 digested with *Bst*EII and *Eco*RI. *Escherichia coli* JK50 transformed with control pBS vector or pCTF-1, -2, -3 constructs were grown at 37 °C and induced with 5 mM IPTG for 1 h. Cells were pelleted and resuspended in SDS sample buffer. Proteins in crude bacterial lysates were fractionated by SDS polyacrylamide gel electrophoresis and transferred to nitrocellulose filters⁴⁸. The CTF-1, -2, -3 products were then visualized by immunoblot analysis using rabbit anti-NF-I/CTF serum² as described previously⁴⁸.

alternative forms of CTF cDNAs consist of different coding regions, their gene products display very similar biochemical properties. First, the site-specific DNA binding activities of products encoded by CTF-1, 2 and -3 are virtually indistinguishable. Each CTF protein recognizes and binds various GCCAAT recognition elements with similar affinity, and the DNase I footprint patterns conferred by the bacterially synthesized CTF proteins are characteristic of the endogenous CTF/NF-I collection of polypeptides isolated from HeLa cells. Moreover, the DNA-binding activity of individual CTF polypeptides do not require the presence of additional factors or other CTF proteins. Second, the product of CTF-1 and -2, when expressed in *Drosophila* cells, potentiate transcription of templates bearing GCCAAT elements with very similar specificities. Finally, the proteins encoded by CTF-1, -2 and -3 are capable of stimulating the initiation of adenovirus DNA replication *in vitro* in a site-dependent way typical of HeLa CTF/NF-I. Thus, these data establish that a single CTF polypeptide can function as both a transcription factor and replication protein.

Why are there multiple CTF proteins? It is interesting to note that the yeast proteins encoded by HAP2 and HAP3 appear to act in concert as site-specific DNA-binding transcription factors²⁴. Moreover, a recent study reports that HeLa cells also contain several CCAAT DNA-binding proteins that require two or more components for efficient binding to occur²⁵. However the data presented here clearly demonstrate that a single CTF polypeptide can bind various CCAAT DNA sequences with high efficiency. Thus, the structural and functional relationship between CTF/NF-I and the CCAAT-binding proteins which apparently require two or more components for function is unclear (see below).

It seems reasonable to postulate that all CCAAT-binding factors form a structurally related family of DNA-binding proteins, which would include the yeast HAP2 and HAP3 proteins because they are apparently functionally homologous to some of the mammalian CCAAT factors²⁶. However, this prediction has not been borne out by a direct comparison of amino-acid sequences between CTF and HAP2 and HAP3. Interestingly, a

search for similarities between CTF and other DNA-binding proteins revealed a limited identity with another multifunctional protein, the yeast RAP1 factor²⁷. A 30 amino-acid domain present in the N terminus of CTF shows 30% identity and 60% similarity with the C-terminal portion of RAP1 (data not shown). Intriguingly, RAP1 has been shown to recognize sequences implicated in both transcriptional repression and activation, as well as in autonomous DNA replication and/or mitotic stability²⁸. Our studies, taken together with other findings, suggest that perhaps the CTF/NF-I factors constitute one structurally distinct family which is part of a larger clan of DNA-binding proteins that recognize CCAAT-related sequences.

Another interesting possibility is that individual members of the CTF family may be differentially regulated. For instance, the alternative protein sequences could contain distinct modification sites allowing differential regulation of individual CTF polypeptides. Alternatively, the various forms of CTF could be selectively expressed in certain tissues or cell types and perhaps at different times during development. Preliminary Northern blot analysis indicates that human teratocarcinoma cells contain a different constellation of CTF mRNA species from that observed in HeLa cells (data not shown). It will be of interest to analyse the pattern of CTF expression in different mammalian cell types as well as during cellular differentiation and embryonic development.

Other CCAAT-binding proteins distinct from CTF/NF-I. There have been a number of recent reports describing the identification of several distinct DNA-binding activities in eukaryotic cells that recognize CCAAT homologies^{10,24,25,29-34}. The best characterized of these other CCAAT-binding factors is a protein designated C/EBP, isolated from rat liver, that binds with high affinity to the recognition sequence GCAAT, rather than GCCAAT^{10,30}. Recently, a partial cDNA clone encoding C/EBP has been isolated³⁵ and DNA-sequence analysis unambiguously established that it is distinct from CTF (unpublished data). In addition to CTF and C/EBP, several other CCAAT-like DNA-binding proteins have been identified^{24,25,29,31-34}. A recent report based on gel-shift assays using HeLa cell extracts suggested the

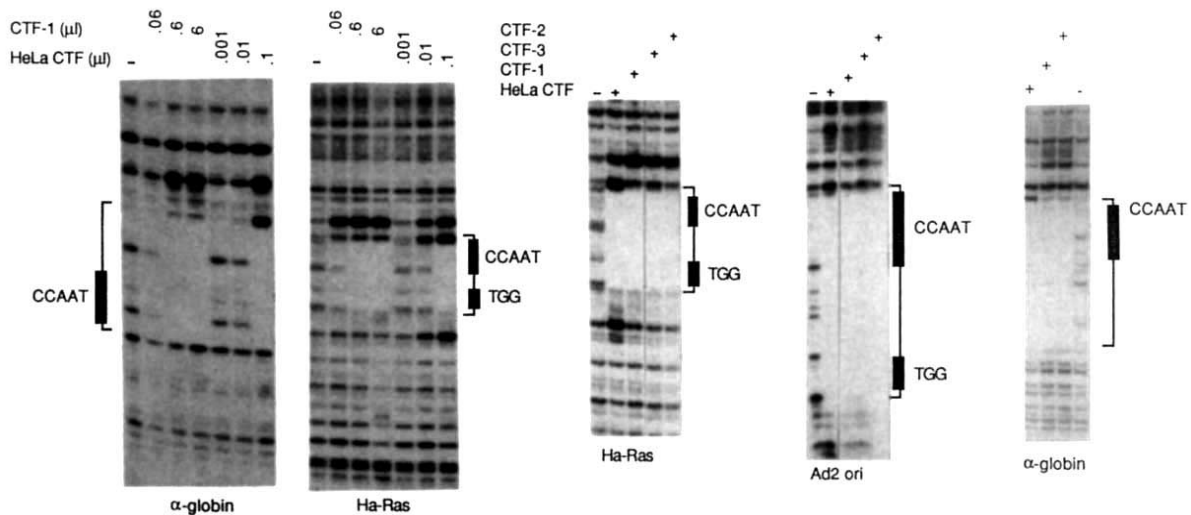


Fig. 4 DNase footprint analysis of affinity purified CTF-1, -2, -3. Templates for DNA-binding assays consisted of the human Harvey ras 1 promoter²¹ (Ha-ras), the Adenovirus type-2 origin of DNA replication¹⁶ (Ad2 ori), and the human α -globin promoter²⁰ (α -globin). Varying amounts of *E. coli* DNA-affinity purified CTF-1, -2 or -3 and HeLa CTF were used in standard DNase footprint reactions². A 3–4-fold higher amount of purified bacterial or HeLa CTF was needed to fully protect the α -globin and Ha-Ras CCAAT boxes, as compared with the Adenovirus sequence (not shown).

Methods. The CTF proteins were extracted from induced *E. coli* cultures containing pCTF-1, pCTF-2 and pCTF-3, according to ref. 49. Bacterial cells were lysed in the presence of lysozyme and sodium deoxycholate. Nucleic acids and associated proteins were precipitated in the presence of 9% polyethylene glycol (PEG) 0.1 M KCl. Proteins were eluted in the presence of 5% PEG, 2 M KCl, diluted to 0.1 M KCl, and loaded onto a DEAE (diethylaminoethyl) DE52 resin column. Flowthrough pooled fractions were added pIdC competitor DNA and loaded on an Adenovirus origin sequence-specific affinity resin². Bound proteins were eluted in the presence of 0.6 M KCl as previously described⁶⁰. About 100 footprint equivalent units were obtained from 500 ml of bacterial culture (1 unit of DNase footprint activity defined as the amount of protein required to bind 1 fmol of the CTF binding site in the Adenovirus origin of DNA replication).

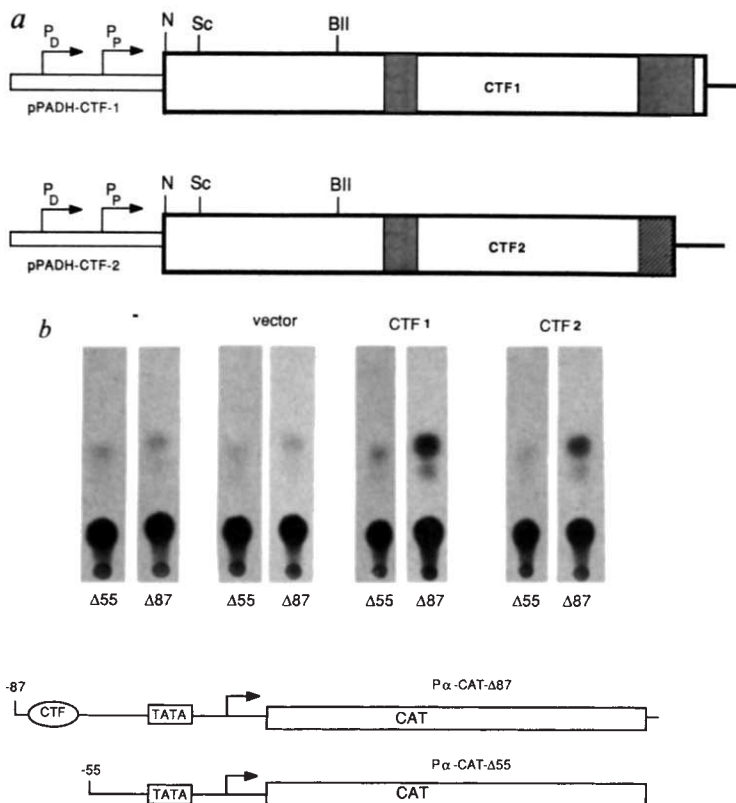


Fig. 5 Stimulation of CCAAT box containing promoters by CTF expressed *in vivo*. **a**, CTF-1 and CTF-2 were expressed from the tandem *Drosophila Adh* promoters in pADH-CTF-1 and pADH-CTF-2. *Drosophila* Schneider cells were co-transfected with CTF-1 or CTF-2 containing expression vectors in presence of the indicated reporter plasmids. **b**, CTF activity was monitored by transactivation of the CCAAT-box containing α -globin promoter, driving transcription of the chloramphenicol acetyl transferase (CAT) gene in α -CAT- Δ 87, or using the fusion of a CCAAT-box truncated promoter to the CAT gene sequence (α -CAT- Δ 55). **Methods.** The pCTF-1 and pCTF-6 *Nco*I-*Xho*I fragment containing CTF coding sequences were inserted in the *Nco*I-*Bam*HI using an *Xho*I linker at the *Bam*HI site, of the pADH expression vector, resulting in the transcriptional fusion of CTF open reading frame to the alcohol dehydrogenase promoter leader sequence. The construction of pADH will be described elsewhere (A. Courey and R.T., manuscript in preparation). The α -globin promoter-CAT fusions were constructed by inserting the mung bean nuclease repaired *Nco*I-*Sca*I fragment of the pSVO α 1p3 deletion mutants⁵¹ into the Klenow repaired *Xba*I site of pBLCAT3 (ref. 52). This results in the insertion of wild-type α -globin sequences from -87 (α -CAT- Δ 87) or from -55 (α -CAT- Δ 55) to position +36, upstream of the CAT gene. Schneider line-2 cell culture and transfection by the Ca-phosphate co-precipitation method were performed by standard methods⁵³. 2 μ g of each expression vectors and/or reporter plasmids were cotransfected, as indicated, in presence of pBSSK⁺ (Stratagene) carrier to a total of 20- μ g DNA. The DNA precipitate was incubated for 32 h in the presence of cells. Cell lysis and CAT assays were performed as previously described⁵³. A 3–5-fold (pADH-CTF-1) or 2–3-fold (pADH-CTF-2) induction of CAT activity was consistently observed in four experiments performed in duplicates, with several independent plasmid preparations.

presence of different CCAAT factors, one (CP1) that binds with high affinity to α -globin CCAAT elements whereas another distinct factor (NF-1), preferentially recognizes the Ad2 origin sites²⁵. It has also been suggested that purified CTF/NFI (ref. 2) represents a mixture of the two distinct CCAAT-binding proteins. But there are several lines of evidence which strongly

argue against this interpretation. First, DNaseI protection and competition footprint assays indicate that HeLa CTF/NF-1 affinity purified over α -globin sequences binds efficiently to the Adenovirus NF-1 site and vice versa². Second, the DNA-binding specificity of purified CTF/NF-1 is not merely a composite of NF-1 and CP1. For example, CTF/NF-1 does not bind to the

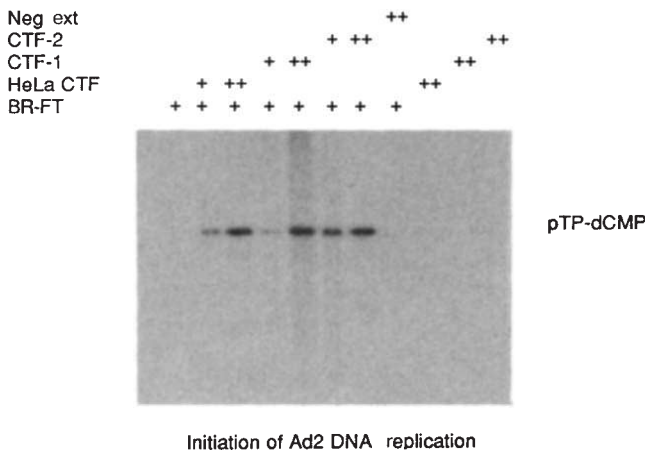


Fig. 6 Stimulation of pTP-dCMP complex formation by CTF. Extracts from HeLa cells and *E. coli* containing pCTF (CTF-1, CTF-2) or control pBS vectors (indicated by neg. ext.) were purified by DNA-affinity chromatography as described in the legends to Figs 1 and 4. The dependence of Adenovirus DNA replication on the presence of CTF was determined by *in vitro* reconstituted experiments as previously described²³. The fraction designated BR-Ft consists of a cellular protein fraction required for optimal initiation of DNA replication (see ref. 23). Preparation of pTP, Ad2 DNA polymerase and viral DNA templates containing the adenovirus origin of DNA replication were a gift of Drs E. O'Neill and T. Kelly. One (+) and two (++) footprint-equivalent units of HeLa or *E. coli* synthesized CTF were used in the assays. The arrow designates the pTP-dCMP initiation complex detected after gel electrophoresis and autoradiography.

CCAAT sequences of MSV-LTR and MHC-2 promoters, unlike CP1. Third, and most compelling, we now demonstrate that purified bacterially synthesized CTF/NF-I represents a single gene product that binds both α -globin and Adenovirus origin sequences efficiently. Most importantly, however, both CTF/NF-I purified to homogeneity from HeLa cells and the individual gene products expressed by cloned CTF *trans*-activate the α -globin promoter and stimulate initiation of adenovirus DNA replication. The discrepancies between DNA binding specificities assigned to the various CCAAT factors could be due, in part, to the various experimental conditions used. It is, perhaps, not surprising that results obtained by DNase footprint assays using purified proteins cannot be directly

compared to gel shift competition assays using crude extracts. Clearly, the purification and determination of specificity by direct DNA-binding assays as well as detailed structural and functional analysis of the different CCAAT-binding proteins will be required to sort out the relationship between these various factors.

How does CTF activate transcription and DNA replication? The finding that a single CTF polypeptide species can serve both as a site-specific transcription factor and initiator of DNA replication suggests an important relationship between replication origins and transcription promoters. In several cases, *cis* control elements important for transcription and replication are located in close proximity, or even juxtaposed, and specific *trans*-acting factors appear to play a dual function governing these activities (refs 36-40; E. O'Neill, C. Fletcher, R. G. Roeder & T. Kelly, manuscript in preparation). The association of transcription factors with origins of replication may provide a simple mechanism for coordinate regulation of transcription and DNA synthesis. In some cases, DNA replication may be dependent on a template that requires activation by RNA polymerase catalysed initiation of transcription. But this situation is apparently not necessary for Adenovirus, SV40 and Polyoma DNA replication, where RNA polymerase II activity is dispensable. Another popular hypothesis envisions the binding of transcription factors to the template as a means of directly altering the chromatin structure to facilitate DNA replication. This seems less likely, at least for adenovirus replication because CTF can act *in vitro* with non-chromatin templates. A more specific model of activation which we favour involves specific interactions between transcription factors bound to the DNA template and components of the replication machinery. Future studies with the CTF cDNA should significantly aid our analysis of specific mechanisms that mediate both replication and transcription.

We thank Drs E. O'Neill and T. Kelly for providing the extracts and templates to assay Adenovirus DNA replication, Dr J. Kadonaga for help and advice on the chromatography techniques, Dr B. Luckow for the gift of plasmids, Dr A. Courey for providing the *Drosophila* expression vectors, D. Lin for help with CAT assays, the members of the laboratory for suggestions and discussions, and K. Ronan for typing the manuscript. This work was supported in part by a grant from the National Institutes of Health to R.T. C.S. was supported in part by an EMBO long-term fellowship, and N.M. by a Swiss NSF Fellowship.

The sequence data in this publication have been submitted to the EMBL/GenBank Libraries.

Received 21 March; accepted 29 April 1988.

- Prujini, G. J. M., Van Driel, W. & van der Vliet, P. C. *Nature* **322**, 656-659 (1986).
- Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J. & Tjian, R. *Cell* **48**, 79-89 (1987).
- Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. *Nucleic Acids Res.* **8**, 127-142 (1980).
- Elstradiadis, A. *et al.* *Cell* **21**, 653-668 (1980).
- McKnight, S. L., Kingsbury, R. C., Spence, A. & Smith, M. *Cell* **37**, 253-262 (1984).
- Siebenlist, U., Hennighausen, L., Battey, J. & Leder, P. *Cell* **37**, 381-391 (1984).
- Hennighausen, L. *et al.* *Nature* **314**, 289-292 (1985).
- Nowock, J., Borgmeyer, U., Püschel, A. W., Rupp, R. A. & Sippel, A. E. *Nucleic Acids Res.* **13**, 2045-2061 (1985).
- Bienz, M. & Pelham, H. R. B. *Cell* **45**, 753-760 (1986).
- Graves, B. J., Johnson, P. F. & McKnight, S. L. *Cell* **44**, 565-572 (1986).
- Theisen, M., Stief, A. & Sippel, A. E. *EMBO J.* **5**, 719-724 (1986).
- Cereghini, S., Raymondjean, M., Carranca, A. G., Herbolme, P. & Yaniv, M. *Cell* **50**, 627-638 (1987).
- Morgan, W. D. *et al.* *Molec. cell. Biol.* **7**, 1129-1138 (1987).
- Jones, K. A., Yamamoto, K. R. & Tjian, R. *Cell* **42**, 559-572 (1985).
- Nagata, K., Guggenheimer, R. A. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6177-6181 (1983).
- Rawlins, D. R., Rosenfeld, P. J., Wides, R. J., Challberg, M. D. & Kelly, T. J., Jr *Cell* **37**, 309-319 (1984).
- Guggenheimer, R. A., Stillman, B. W., Nagata, K., Tamanoi, F. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3069-3073 (1984).
- Leegwater, P. A. J., Van Driel, W. & van der Vliet, P. C. *EMBO J.* **4**, 1515-1521 (1985).
- Rosenfeld, P. J. & Kelly, T. J. *J. biol. Chem.* **261**, 1398-1408 (1986).
- Liebhaber, S. A., Gossens, M. J. & Kan, Y. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7054-7058 (1980).
- Ishii, S., Merlino, G. T. & Pastan, I. *Science* **230**, 1378-1381 (1985).
- Benayahu, C., Spoerel, N., Haymerle, H. & Ashburner, M. *Cell* **33**, 125-133 (1983).
- Rosenfeld, P. J., O'Neill, E. A., Wides, R. J. & Kelly, T. J. *Molec. cell. Biol.* **7**, 875-886 (1987).
- Olesen, J., Hahn, S. & Guarente, L. *Cell* **51**, 953-961 (1987).
- Chodosh, L. A., Baldwin, A. S., Carthew, R. W. & Sharp, P. A. *Cell* **53**, 11-24 (1988).
- Chodosh, L. A. *et al.* *Cell* **53**, 25-35 (1988).
- Shore, D. & Nasmyth, K. *Cell* **51**, 721-732 (1987).
- Brand, A. H., Micklem, G. & Nasmyth, K. *Cell* **51**, 709-719 (1987).
- Dorn, A., Bollekens, J., Staub, A., Benoist, C. & Mathis, D. *Cell* **50**, 863-872 (1987).
- Johnson, P. F., Landschulz, W. H., Graves, B. J. & McKnight, S. L. *Genes Dev.* **1**, 133-146 (1987).
- Lichtsteiner, S., Wuarin, J. & Schibler, U. *Cell* **50**, 963-973 (1987).
- Miksicsek, R., Borgmeyer, U. & Nowock, J. *EMBO J.* **6**, 1355-1360 (1987).
- Oikarinen, J., Hatamochi, A. & deCrombrughe, B. *J. biol. Chem.* **262**, 11064-11070 (1987).
- Raymondjean, M., Cereghini, S. & Yaniv, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 757-761 (1988).
- Landschulz, B. & McKnight, S. *Genes Dev.* (in the press).
- Salzman, N. P., Venkatachala, N. & Selzer, G. B. in *The Papovaviridae* Vol. 1. *The Polyomaviruses* (ed. Salzman, N. P.) 27-98 (Plenum, New York, 1986).
- Veldman, G. M., Lupton, S. & Kamen, R. *Molec. cell. Biol.* **5**, 649-658 (1987).
- Stenlund, A., Bream, G. L. & Botchan, M. R. *Science* **236**, 1666-1671 (1985).
- Reisman, D. & Sugden, B. *Molec. cell. Biol.* **6**, 3838-3846 (1988).
- DePamphilis, M. L. *Cell* **52**, 635-638 (1988).
- Rüegg, U. T. & Rüdiger, J. in *Methods in Enzymology* Vol. 47 (eds Hirs, C. H. W. & Timasheff, S. N.) 111-116 (Academic, New York, 1977).
- Andrews, P. C. & Dixon, J. E. *Anal. Biochem.* **161**, 524-528 (1987).
- Lathe, R. *J. molec. Biol.* **183**, 1-12 (1985).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Hauser, C. A. *et al.* *Cell* **43**, 19-28 (1985).
- Santoro, C. *et al.* *Nucleic Acids Res.* **14**, 2863-2876 (1986).
- Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. *Cell* **51**, 1079-1090 (1988).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
- Gross, C., Engbaek, F., Flamang, T. & Burgess, R. J. *Bact.* **128**, 382-389 (1976).
- Kadonaga, J. T. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5889-5893 (1986).
- Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. *Cell* **27**, 279-288 (1981).
- Luckow, B. & Schütz, G. *Nucleic Acids Res.* **15**, 5490 (1987).
- DiNocera, P. D. & Dawid, I. B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7095-7098 (1983).

Nature and evolution of the eclipsing millisecond binary pulsar PSR1957+20

W. Kluźniak, M. Ruderman, J. Shaham & M. Tavani

Physics Department and Astrophysics Laboratory, Columbia University, New York, New York 10027, USA

A millisecond pulsar may be able to evaporate a very light companion by a particular component of its energetic radiation¹⁻³. Pulsar turn-on in the binary follows a stage of mass transfer in a very low-mass X-ray binary (LMXB), dominated by an evaporative wind from the surface of a very light, at least partially degenerate companion¹. The wind is driven by a large MeV γ -ray flux powered by an accretion dynamo. That source of irradiation ceases when it is replaced by that from the millisecond pulsar, which has been spun up by accretion. We apply this model to the recently discovered⁴ 1.6-ms pulsar PSR1957+20.

The wind may carry mass away from the companion at a rate of 10^{15} – 10^{16} g s⁻¹. In several hundred million years, the companion could evaporate, leaving behind a solitary millisecond pulsar similar to PSR1937+214. At present the inferred radius of the companion is likely to be less than its Roche lobe. The ratio of the radius of a degenerate companion to its Roche lobe can decrease as long as more than one third of the mass lost by the companion is transferred to an accretion disk; this is achieved in the LMXB stage when the secondary's wind velocity drops below the escape velocity from the system. This ratio of radii stops decreasing abruptly when the pulsar turns on and its kHz dipole radiation expels the accretion disk. Although part of the energetic radiation from the pulsar can still sustain a very strong wind from the secondary, the pulsar's dipole radiation sweeps it out of the system. The eclipse duration of PSR1957+20, at 430 MHz, implies an eclipsing region much bigger than the $0.02M_{\odot}$ secondary's Roche lobe. This provides the first experimental support, in a binary system, for an ionized wind.

Of the various kinds of incident radiation which can drive a wind from a secondary, MeV radiation is the most effective², because the wind will be formed primarily in that region of the stellar atmosphere where the radiative cooling timescale exceeds the dynamical expansion timescale. Gamma-rays of MeV energy, unlike hard X-rays, deposit most of their energy in a single scattering and, unlike soft X-rays, have a cross-section which is not diminished in a highly ionized outer atmosphere². For these reasons, MeV radiation deposits relatively more energy in the relevant (outer) layers of the stellar atmosphere than other kinds of illumination. The evaporative mass-loss rate, $\dot{m}_w$, is²

$$\dot{m}_w \approx -3 \times 10^{-17} \xi \frac{\hat{L}}{\hat{\sigma}} \text{ g s}^{-1} \quad (1)$$

where $\hat{L}$ is the total intercepted MeV γ -ray radiation power, $\hat{\sigma}$ is the column density at which the radiation is absorbed ($\hat{\sigma} \approx 3 \text{ g cm}^{-2}$ at MeV energies), and ξ is defined below. If the MeV radiation leaving the primary and its immediate environment is equivalent to a γ -ray luminosity L_{γ} , then $\hat{L} = L_{\gamma}(R_c/2a)^2$ where R_c is the radius of the companion and a is the binary separation. The evaporative wind velocity is² $v_w = (L_{\gamma}h/4\pi\hat{\sigma}a^2)^{1/2}$ where $h \leq R_c$ is the scale height of the companion's atmosphere. If $v_8 \equiv v_w/(7 \times 10^7 \text{ cm s}^{-1}) \geq 1$, $\xi = 1$; for $0.1 \leq v_8 \leq 1$, $\xi = v_8^2$. The efficiency of conversion of MeV radiation to kinetic energy of a wind is $f = 0.025\xi v_8^2$.

The inferred existence of a wind in PSR1957+20 (which is known not to be a strong X-ray source) suggests that the companion is irradiated by energetic ($\sim$ MeV) γ -rays. The inference of previous evolution without Roche lobe overflow implies the probable existence of such incident radiation in the past. We have discussed elsewhere^{2,5} the expected MeV radiation directly from the vicinity of the spun-up neutron star in an accretion

LMXB. For a companion of a non-accreting ms radio pulsar, the emitted high-energy radiation from near the neutron star may be expected to be mainly in a TeV e^+ wind which would produce synchrotron MeV γ -rays near the secondary. We assume that the magnetic dipole moment (and hence the spin-down rate) of PSR1957+20 is similar to that of PSR1937+21 and other ms pulsars, implying a surface magnetic field $B_s \approx 5 \times 10^8$ G. The pulsar's spin-down power is $L_p \approx 5 \times 10^{36} \text{ erg s}^{-1}$, and its lifetime is expected to be $\tau \approx 10^8$ yr.

The evolutionary scheme we describe can lead to systems like PSR1957+20. We distinguish four phases according to the origin and characteristics of the secondary's mass loss. In phase I, for companion mass $m > 0.1M_{\odot}$, the system evolves by standard Roche lobe overflow, driven by gravitational radiation. In phase II, Roche lobe overflow close to the Eddington limit is maintained by some other mechanism, perhaps evaporative mass loss from the companion due to MeV illumination from the magnetosphere-disk boundary region of the primary. In phase III, which is central to our model of how the system came into its present state, conservative evolution (no angular momentum loss from the system) proceeds through mass transfer sustained by an evaporative wind feeding an accretion disk. At the close of phase III, the pulsar turns on and expels the accretion disk, and the system enters phase IV, the present phase of pulsar-driven evaporation of the secondary. During phase III, the neutron star slowed from ~ 1 ms to its present period, the mass of the secondary dropped from $\sim 0.04M_{\odot}$ to the current value of $0.02M_{\odot}$, and the binary period increased from ~ 1 h to 9 h.

If we assume, in the mass range of interest, that a degenerate dwarf has a radius $R_c \propto m^{-1/3}$, then $R_L/R_c \propto am^{2/3}$, where R_L is the Roche lobe radius. (This approximation¹³ is adequate here; the rates of heating and mass loss in our model are low enough to justify the assumption that the companion star is almost degenerate and near thermal equilibrium.) Quasistatic Roche lobe overflow by a degenerate secondary then implies that the orbital separation $a \propto m^{-2/3}$; on the other hand, the fundamental equation of binary evolution with small secondary mass, $\delta a/2a + \delta m/m = \delta J/J$, yields $a \propto m^{-2}$ whenever the total angular momentum J of the system is constant. This limiting case corresponds to the largest possible increase of a/R_c and R_L/R_c (per unit mass loss) because, in general, $\delta J/J \leq 0$. More generally, the assumption that $\delta J/J = k(\delta m)/m$ yields $a \propto m^{-2(1-k)}$ whenever k is a constant. From the definitions of R_L and R_c one obtains the important condition

$$-\frac{d}{dm} \ln \left(\frac{R_L}{R_c} \right) = \left(\frac{2}{3} - k \right) \frac{2}{m} \quad (2)$$

Thus, in the phase when R_L/R_c is constant, as happens during Roche lobe overflow, $k = 2/3$. R_L/R_c increases when $2/3 > k \geq 0$. In those phases when the total angular momentum is conserved ($k = 0$), $R_L/R_c \propto m^{-4/3}$. The physical significance of equation (2) becomes clearer if the mass-loss rate of the companion, $\dot{m}$, is divided into two parts, the mass ejected from the system, at the rate $\dot{m}_s \leq 0$, and the neutron star disk accretion rate $\dot{M} \geq 0$, so that $-\dot{m}_s + \dot{M} = -\dot{m}$. The accretion disk recycles angular momentum, transporting it outward through viscous torques and eventually returning it through tidal interactions to the secondary.

In the late stages of evolution of the progenitor LMXB, orbital angular momentum loss due to gravitational radiation is negligible on the timescales of interest, so essentially the whole angular momentum loss may be due to mass ejection (with or without magnetic braking):

$$\frac{\dot{J}}{J} = \beta \frac{\dot{m}_s}{m} = \beta \frac{\dot{M} + \dot{m}}{m} \quad (3)$$

where $\beta - 1$ is a measure of the angular momentum gained by the wind as it leaves the system owing to tidal interactions or magnetic braking. From equation (2), during late evolution of

the binary system, $R_L/R_c = \text{const}$ iff $\dot{M} = (1 - 2/(3\beta))|\dot{m}|$, while R_L/R_c increases iff $|\dot{m}| > \dot{M} > [1 - 2/(3\beta)]|\dot{m}|$.

We argue that in phase III, directly before the current stage of evolution of the PSR1957+20 system, and thus before the kHz pulsar dipole radiation turned on, the primary had a disk such that the accretion power L_a was less than the spin-down power $L_p \approx 5 \times 10^{36} \text{ erg s}^{-1}$. The system was an LMXB with low ($\approx L_p$) X-ray luminosity. Dissipation of the rotational energy of the neutron star in the inner region of the accretion disk may yield an effective γ -ray luminosity L_γ larger than L_a corresponding to the same accretion rate¹. We assume $L_\gamma \approx L_p$ in this phase. The evaporative wind velocity $v_w < v_{\text{esc}}$, the escape velocity from the binary system. During this phase the simplest model of conservative evolution applies, with $\dot{m}_s = 0$. Because $\delta J = 0$, the system undergoes rapid evolution (maximum possible $(|da/dm|)$ with the ratio R_L/R_c increasing from 1 to the current value, ~ 3 , corresponding to a decrease of m by a factor ~ 2 and to an increase of the binary period P_{orb} from about 1 h to the presently observed 9.17 h ($P_{\text{orb}} \propto m^{-3}$). Because the change of the rotational energy of the neutron star E_Ω satisfies $\Delta E_\Omega = \int L_p dt \propto \int \dot{m} dt = \Delta m$ with the constant of proportionality obtained from equation (1), $L_p = L_\gamma$ implies $\Delta E_\Omega \approx 4 \times 10^{51} \text{ erg}$, a value comparable to the present rotational energy of PSR 1957+20.

This raises the question of how the system entered this phase with the neutron star spinning with a millisecond period. The early evolution could have been driven by gravitational radiation according to the canonical LMXB scenario. To avoid substantial spin-down of the neutron star⁶ a model for phase II has been proposed² in which winds sustained by γ -rays from the secondary drive the angular momentum loss from the system, keeping $\dot{M}$ near the Eddington limit even when $m < 0.1 M_\odot$. Mass transfer within the system unavoidably undergoes a qualitative change when the companion mass drops below a certain critical value m_c , which is several hundredths of a solar mass. Before this transition, L_γ is insensitive to m and has the highest value possible for the system, close to the Eddington value. At the transition, $\dot{M}$ abruptly decreases and L_γ drops by a factor ≈ 10 , approaching L_p . Subsequent evolution (phase III) proceeds with $L_\gamma \approx \text{constant}$, decreasing $\dot{m} = \dot{M}$, and J constant.

As the rate of the mass transfer into the accretion disk drops with increasing binary separation, the accretion disk inner radius r_a moves toward the neutron star light cylinder radius ($c/\Omega = 5 \times 10^6 \text{ cm}$ ($P_s/10^{-3} \text{ s}$), where P_s is the rotational period of the neutron star). The radius r_a has been estimated¹ to pass beyond the light cylinder radius when $|\dot{M}| \leq 10^{16} (L_p/10^{37} \text{ erg s}^{-1}) \text{ g s}^{-1}$. In this regime the magnetic dipole radiation pressure of a non-aligned radiopulsar can push away ionized accreting matter and prevent the formation of an accretion disk. Matter evaporated from the companion will now be swept out the system and there can no longer exist an accretion dynamo to generate MeV γ -rays. In the resulting near vacuum around it, the rapidly rotating spinning neutron star behaves as an isolated ms pulsar. The fractional mass-loss rate of the companion due to any wind will be approximately equal to the fractional rate of loss of angular momentum of the system ($\beta \delta m/m = \delta J/J$). The companion's wind results in no further large increase in a or the binary period.

In phase IV, the energetic radiation from a spinning-down solitary radiopulsar will be qualitatively very different from the putative MeV luminosity of an accreting LMXB. Models for the Crab nebula suggest^{7,8} that the pulsar spin-down energy loss is mainly into a $10^{39} \text{ s}^{-1} e^\pm$ flux of TeV energy. An analysis of X-ray synchrotron halos around radio pulsars⁹ with spin-down power $\geq 10^{34} \text{ erg s}^{-1}$ leads to the expectation of a similar flux for all such strong field radiopulsars. The production of such huge pair fluxes is a natural consequence of outer magnetosphere models for Crab- and Vela-like radiopulsars¹⁰. Subsequent acceleration to TeV energies would come from the interaction of this $e^\pm$ plasma with the enormously strong pulsar magnetic

dipole radiation near the star^{11,12}. Synchrotron radiation by the TeV $e^\pm$ pairs would occur only if the TeV pairs meet a magnetic field other than the coherent ordered one which initially accelerated them.

The whole magnetosphere of a millisecond pulsar resembles the outer magnetosphere of the Crab or Vela pulsars in current flow, potential drops and field strength. If the same fraction of its spin-down power were radiated into MeV γ -rays as in the Crab pulsar ($< 10^{-3}$), the resulting mass loss, $|\dot{m}| \leq 4 \times 10^{12} \text{ g s}^{-1}$, may or may not be too small to account for the observed secondary wind. On the other hand, a flux of TeV $e^\pm$ qualitatively similar to that from the Crab pulsar could be even more effective in sustaining the wind. Because a solitary ms pulsar has no supernova remnant to confine the magnetic field, X-ray synchrotron halos need not exist around them. However, a companion can produce a magnetic field irregularity of order 100 G to give $(E_\pm/m_e c)^2 (e\hbar B/m_e c) \approx \text{MeV}$ synchrotron radiation in its neighbourhood from the TeV $e^\pm$ flux to which it is exposed. ($E_\pm$ is the energy of the $e^\pm$ pairs.) This irregularity can be either the $(L_p/a^2 c)^{1/2} < 10^2 \text{ G}$ field of the reflected kHz wave, or a magnetic field of the degenerate dwarf companion with surface strength $B > 10^2 \text{ G}$. In either case we predict that MeV γ -rays of luminosity $\approx 10^{33} (B/100 \text{ G})^{2/3} \text{ erg s}^{-1}$ are emitted from the system, modulated at the orbital period, 9.17 h, with maxima near orbital phase 0 (radio eclipse) and minima at phase 0.5. This orbital correlation is opposite to that expected from optical emission caused by the total illumination ($\approx L_\odot$) of the secondary by all kinds of energetic radiation which reach its surface. A TeV $e^\pm$ flux $\approx (1/2)L_p \approx 3 \times 10^{36} \text{ erg s}^{-1}$ would then give at the companion $\dot{L} \approx 6 \times 10^{32} \text{ erg s}^{-1}$ in MeV γ -rays and, from equation (1), $\dot{m}_w \approx 5 \times 10^{15} \text{ g s}^{-1}$. This would be sufficient to evaporate PSR1957+20's companion and leave it as an isolated ms pulsar before it spins down. (But direct MeV γ -rays expected from the ms pulsar may be sufficient to sustain the wind if $|\dot{m}_w| \leq 4 \times 10^{12} \text{ g s}^{-1}$; if such a weak wind is maintained the companion would then not be expected to evaporate during the pulsar spin-down lifetime.)

The eclipse data of PSR1957+20 indicate a complicated wind geometry⁴. The post-eclipse delays are ~ 10 times larger than the pre-eclipse ones, and disappear over a phase range comparable to the eclipse width, ≈ 0.1 , but the eclipse is symmetrical. Future observations will tell whether the eclipsing region is limited by the wind being swept away by the kHz radiation from the pulsar or is defined by some other mechanism that keeps the density profile quasi-spherical close to the companion. In the first case, the time delay data suggest a comet-like wind tail. Close to the companion its shape is a truncated cone whose axis lies radially away along the pulsar-companion line; it bends backwards only at large distances. The dispersive delay δt on the egress side of the eclipse may then be mainly a geometrical effect due to a long line of sight through the free electrons in the 'comet tail' far away from the companion where plasma frequency lies far below the frequency of the pulsar microwave beam. Even though these time delays may be from a low-density plasma, the main eclipse can still be caused by overdense plasma. A consistent set of parameters that fits the above model for strong MeV γ -ray irradiation of the secondary is $B \approx 20 \text{ G}$, $v_w \approx 3 \times 10^7 \text{ cm s}^{-1}$, $E_\pm \approx 1 \text{ TeV}$, $\dot{m}_w = 5 \times 10^{15} \text{ g s}^{-1}$. These give a bow shock between wind and pulsar dipole field at a distance of $\approx 5 \times 10^{10} \text{ cm}$ from the secondary (roughly the 430-MHz eclipse radius) around which the plasma is overdense for frequencies $\leq 400 \text{ MHz}$.

We thank D. Stinebring and J. Taylor for kindly sharing results on PSR1957+20 prior to publication. B. Paczyński and A. Fruchter are thanked for helpful conversations. We acknowledge partial support of this research through NASA and NSF grants.

Received 7 April; accepted 8 June 1988.

1. Ruderman, M., Shoham, J. & Tavani, M. *Astrophys. J.* (in the press).

2. Ruderman, M., Shoham, J., Tavani, M. & Eichler, D. *Astrophys. J.* (in the press).

3. Kluźniak, W., Ruderman, M., Shaham, J. & Tavani, M. in *Proc. Conf. on the Physics of Neutron Stars and Black Holes* (eds Tanaka, Y. & Koyama, K.) (in the press).
4. Fruchter, A. S., Stinebring, D. R. & Taylor, J. H. *Nature* **333**, 227–239 (1988).
5. Kluźniak, W., Ruderman, M., Shaham, J. & Tavani, M. *Nature* (submitted).
6. Jeffrey, L. C. *Nature* **319**, 384–386 (1986).
7. Kennel, C. F. & Coroniti, F. V. *Astrophys. J.* **283**, 710–730 (1984).
8. Rees, M. J. & Gunn, J. E. *Mon. Not. R. astr. Soc.* **167**, 1–12 (1974).
9. Helfand, D. *Adv. space Res.* **3**, 29–34 (1984).
10. Cheng, K. S., Ho, C. & Ruderman, M. *Astrophys. J.* **300**, 522–539 (1986).
11. Rees, M. J. *Nature* **230**, 55–57 (1971).
12. Gunn, J. E. & Ostriker, J. P. *Astrophys. J.* **165**, 523–541 (1971).
13. Rappaport, S., Nelson, L. A., Ma, C. P. & Joss, P. C. *Astrophys. J.* **322**, 842–851 (1987).

Fate of the companion stars of ultra-rapid pulsars

E. P. J. van den Heuvel & J. van Paradijs

Astronomical Institute “Anton Pannekoek”,
University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam,
The Netherlands

A millisecond pulsar that is formed by spin-up ‘recycling’^{1,2} in a binary system will, once the mass transfer becomes temporarily interrupted, start to evaporate its companion star as a consequence of the large impinging pulsar energy flux. This evaporation is easiest if the pulsar has a short pulse period, the companion star has a relatively large radius and is therefore hydrogen-rich, and the orbital period is short. Evaporation of companion stars induced by millisecond pulsars could account for the lack of low-mass X-ray binaries with short orbital periods below the period gap of the cataclysmic variables, and for the statistics of new-born radio pulsars and their space velocities.

Of the two millisecond pulsars with the shortest pulse periods, both ~ 1.6 ms, one is single³, and the other has a companion that appears to be evaporating⁴. This is evidenced by its very low mass of $\sim 0.02 M_{\odot}$, and its 50-min eclipse duration (per orbital cycle of 9 h), which indicates that the companion is surrounded by a ‘halo’ extending to almost a factor of three outside its Roche lobe. Were this halo a stationary atmosphere, it would immediately flow away and be lost. Therefore it must be a dynamic phenomenon, namely a stellar wind that is continuously replenished by the secondary.

We suggest that this wind arises from the large energy flux from the pulsar impinging on the stellar surface. Ruderman *et al.*^{5,6} recently discussed possible mechanisms that could drive such a wind, and proposed that MeV γ -rays, possibly produced through synchrotron radiation of TeV $e^{\pm}$ pairs in a 10^2 gauss magnetic field near the companion star surface, would provide a very effective mechanism.

The 1.55 ms pulsar PSR1937+21 (ref. 3), despite its weak surface magnetic field strength of 4×10^8 G (as inferred from its period derivative⁷), is a very bright pulsar having a total rotational energy loss rate $dE_{\text{rot}}/dt = 4\pi^2 I \dot{P} \sim 10^{36}$ erg s⁻¹ (we have assumed the standard neutron star moment of inertia $I = 10^{45}$ g cm²; $\dot{P} = 10^{-19}$, $P = 1.55$ ms). This rate can be expressed in terms of the pulsar’s surface dipole magnetic field strength B_s , and rotational period P , as⁸ $L_p = dE_{\text{rot}}/dt = (2R_n^6/3c^3)B_s^2(2\pi/P)^4$ where R_n is the neutron star radius, and c the speed of light.

Assuming to a first approximation that the energy flux is radiated in a spherically symmetric way, a fraction $(R/2a)^2$ of L_p will impinge on the companion star with radius R at a distance a . We make the simple assumption that the fraction f of the impinging flux is used to drive a wind from this companion (depending on the assumptions made about the detailed physics, the results of Ruderman *et al.*^{5,6} show that f can easily be 0.1 and may even approach unity). If we assume that this wind has a characteristic outflow velocity equal to the escape velocity, v_e , from the companion (as stellar winds tend to⁹), we have the

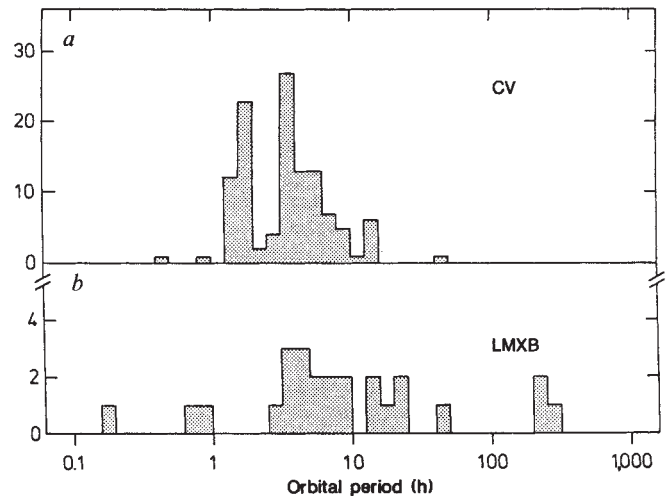


Fig. 1 Orbital period distributions for *a*, the cataclysmic variables and *b*, low-mass X-ray binaries. Data for the cataclysmic variables have been taken from Ritter¹¹. Data for the low-mass binaries are from Parmar and White¹², except that we have not included LMC X-2 (for which recent optical observations give no evidence for a 6.4 h period), and have included the 15.1 h period system Cen X-4 (Chevalier, C., Ilovaisky, S., Van Paradijs, J., Van der Klis, M. & Pedersen, H., unpublished results).

following equation for the rate $\dot{M}$ of mass loss: $\frac{1}{2}\dot{M}v_e^2 = f(R/2a)^2(2R_n^6/3c^3)B_s^2(2\pi/P)^4$. Using $v_e^2 = 2GM/R$, the evaporation timescale τ_{evap} of the companion star becomes of order

$$\tau_{\text{evap}} = M/\dot{M} = (3c^3/2R_n^6)(2GM^2/R^3)(a^2/16\pi^4 f)P^4/B_s^2 \quad (1)$$

On the other hand, the spin-down timescale of the pulsar is

$$\tau_{\text{sd}} = P/\dot{P} = (P^2/B_s^2)(3c^3/8\pi^2 R_n^6)I \quad (2)$$

Combining equations (1) and (2), we obtain the simple expression

$$\tau_{\text{evap}}/\tau_{\text{sd}} = (GM^2/R)(a/R)^2(P^2/2\pi^2 fI) \quad (3)$$

In these expressions R itself is, of course, a function of M . For a non-degenerate main-sequence star we have approximately¹⁰

$$R/R_{\odot} = M/M_{\odot} \quad (4)$$

and τ_{evap} scales as M^{-1} . For low values of M the star is degenerate, and¹⁰

$$R/R_{\odot} = 0.013(1+X)^{5/3}(M/M_{\odot})^{-1/3} \quad (5)$$

where X is the hydrogen mass fraction. Then τ_{evap} scales as M^3 . This shows that for very low companion masses the relative evaporation timescale becomes very short. For $X = 0.70$, the relations (4) and (5) fit to one another at $M = 0.075 M_{\odot}$.

It is of interest to note that equation (3) is independent of the magnetic field strength of the neutron star. Expressing masses and radii in solar units, and P in the value for the 1.55 ms pulsar, we obtain

$$\tau_{\text{evap}}/\tau_{\text{sd}} = [10^{-2}/(f/0.1)](a/R)^2(P/1.55 \text{ msec})^2(M/M_{\odot})^2/(R/R_{\odot}) \quad (6)$$

This equation shows that a companion star like the Sun, at a distance of $a = 3R$, will be completely evaporated for $f = 0.1$ by a pulsar companion if the pulse period is less than 5 ms. If the companion mass is only $0.3 M_{\odot}$ (as is the case at the upper boundary of the ‘period gap’, see below), even a 10 ms pulsar will evaporate it for $f = 0.1$, and a 3 ms pulsar will do so for $f = 0.01$.

For the case of the 1.6 ms binary pulsar, with $M = 0.02 M_{\odot}$, $a = 2.5 R_{\odot}$, and $R = 0.166 R_{\odot}$ (according to equation 5), we obtain $\tau_{\text{evap}}/\tau_{\text{sd}} = 1.6 \times 10^{-2}$, for $f = 0.1$. If we assume that this pulsar has the same spin-down timescale (2×10^8 yr) as the

1.55 ms pulsar, then the $0.02 M_{\odot}$ companion of the 1.6 ms pulsar will disappear in the next ten million years.

In the past evaporation must have been much slower. If we assume that the orbital separation has not changed much (for the case of more or less radial outflow, this is plausible), then the maximum value of the evaporation timescale was $\sim 10^8$ yr (when the companion became degenerate at a mass of $0.075 M_{\odot}$). This maximum timescale is roughly the total timescale for evaporation of the companion.

If the pulsar-wind model described above for PSR1957+20 is correct, it may be relevant in explaining the difference in the orbital-period distributions of low-mass X-ray binaries (LMXB) and cataclysmic variables (CVs). These two types of mass-exchanging low-mass binaries differ only in the nature of their compact mass-accreting component, namely, a neutron star in a LMXB and a white dwarf in a CV. From Fig. 1, in which these orbital-period distributions are plotted, it appears that there are no LMXB with periods below the period gap (between 2 and 3 h) in the distribution of the CVs. Below this gap the period distribution of the CVs shows a peak between 80 min and 2 h. We only find three LMXB with very short periods below the 80-min minimum period for CVs (ref. 13). We note that, since the extensive X-ray observations of LMXB with EXOSAT (which were not interrupted every ~ 100 min by Earth occultations, as they were in previous X-ray satellites), there is no obvious selection effect against the detection of orbital periods of LMXB that would operate in the 1 to 2 h range but not in the 3 to 4 h range, say; the same applies to optical observations, which can generally be carried out during a number of consecutive nights for time intervals of at least 6 h per night.

Mass transfer in the CVs and LMXB for orbital periods below about 10 h is driven by loss of orbital angular momentum, which is due to gravitational radiation and is supplemented (for orbital periods above ~ 3 h) by magnetic braking (see ref. 14 for example). Thus, the evolution of these systems is governed by a secular decrease of the orbital period. The gap in the period distribution of CVs between 2 and 3 h could be related to the cessation of magnetic braking¹⁴ when the orbital period has decreased to 3 h (corresponding to a companion mass of about $0.30 M_{\odot}$). This would cause a drop in the mass transfer rate. The companion star, which was somewhat out of thermal equilibrium and therefore slightly puffed up, will shrink below its Roche lobe when the mass transfer decreases. Mass transfer then stops completely until the orbital separation has decreased sufficiently as a result of gravitational radiation, when the secondary again fills its Roche lobe (this happens at a period of about 2 h).

The above discussion shows that when mass transfer stops at the upper edge of the period gap (near 3 h) the LMXB contain an important source of power that is absent in the CVs—a rapidly rotating magnetized neutron star which accompanies the low-mass secondary. An estimate of the expected rotation frequency of the neutron star is given by the equilibrium frequency for a magnetized neutron star accreting from an accretion disk. The exact value of this equilibrium frequency ν_{eq} depends on the details of the accretion model; using the accretion model of Ghosh and Lamb¹⁵, we find for accreting neutron stars (with canonical values for its mass and radius of $1.4 M_{\odot}$ and 10 km)

$$\nu_{\text{eq}} = (0.25 \text{ Hz}) L_{37}^{3/7} \mu_{30}^{-6/7} \quad (7)$$

Here L_{37} is the X-ray luminosity in units of $10^{37} \text{ erg s}^{-1}$, and μ_{30} is the magnetic moment of the neutron star. With typical values for LMXB with periods near a few hours of $L_{37} \sim 1$, an assumed $\mu_{30} = 10^{-3.5}$, corresponding to the bottom field for old

neutron stars^{16,17}, we find spin frequencies in the range 250–500 Hz.

From equations (6) and (7) we find that for a system with a period at the upper edge of the period gap (secondary mass $\sim 0.30 M_{\odot}$), the ratio $\tau_{\text{evap}}/\tau_{\text{sd}} < 1$, so spin frequencies in this range are sufficient to evaporate the secondary. The absolute timescale for evaporation (assuming $f = 0.1$) is comparable to the timescale needed to cross the period gap by the emission of gravitational radiation.

If the secondary star has a sufficiently developed helium core, it is possible (depending on the exact value of the neutron star spin period) that this core survives the evaporation process (because of its relatively small size, and corresponding high escape velocity). Such a system will evolve with a decreasing orbital period (due to gravitational radiation) until it comes into contact again at a period much shorter than 2 h. This provides a natural explanation for the occurrence of LMXB (such as the 50-min system 4U 1916–05) below the (~ 80 -min) minimum period for CVs.

We finally note that the above described evaporation process may have a major effect on (1) estimates of the birth rates of low-mass X-ray binaries through accretion-induced collapse of a white dwarf in a CV (refs 18 and 19), and (2) the interpretation of the statistics of new-born radio-pulsars and their space velocities. A large fraction of all pulsars are expected to be born in very close post-common-envelope evolution binaries with a main-sequence companion of relatively low mass (at most a few solar masses)²⁰. The exploding stars in such systems are the bare helium cores of stars initially more massive than $\sim 8 M_{\odot}$. The supernova mass ejection imparts a large runaway velocity to such systems ($100\text{--}300 \text{ km s}^{-1}$), but in general does not disrupt them²⁰. If pulsars are born with periods of ~ 10 ms or less, and with strong surface magnetic fields ($> 10^{12} \text{ G}$), then they will evaporate their companions on timescales ranging from a few years to perhaps a few 10^4 yr (see equation (2)) and become single runaway pulsars. Such a model would at once explain why so many pulsars have large runaway velocities, and why so strikingly few pulsars are found in binaries. In fact, this striking absence of close companions might, in reverse, be taken as evidence for the fact that in supernovae pulsars are formed with ms periods and strong magnetic fields.

We thank Dr W. H. G. Lewin for bringing the discovery of the 1.6 ms binary pulsar to our attention, Drs J. H. Taylor and A. S. Fruchter for communicating their early results, and Dr M. S. Ruderman for discussion.

Received 28 March; accepted 16 June 1988.

- Alpar, M. A., Cheng, A. F., Ruderman, M. A. & Shaham, J. *Nature* **300**, 728–730 (1982).
- Radhakrishnan, V. & Srinivasan, G. *Proc. Second Asian-Pacific Regional Meeting on Astronomy* (eds Hidayat, B. & Feast, M. W.) 423–431 (Tira Pustaka, Jakarta, 1984).
- Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M. & Goss, W. M. *Nature* **300**, 615–618 (1982).
- Fruchter, A. S., Stinebring, D. R. & Taylor, J. H. *Nature* **333**, 297–299 (1988).
- Ruderman, M. A., Shaham, J., Tavani, M. & Eichler, D. *Astrophys. J.* (in the press).
- Ruderman, M. A., Shaham, J. & Tavani, M. *Astrophys. J.* (in the press).
- Backer, D. C., Kulkarni, S. R. & Taylor, J. H. *Nature* **301**, 314–315 (1983).
- Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
- Abbott, D. C. *Astrophys. J.* **225**, 893–901 (1978).
- Savonije, G. J. in *Accretion-driven Stellar X-ray Sources* (eds Lewin, W. H. G. & van den Heuvel, E. P. J.) 343–365 (Cambridge University Press, 1983).
- Ritter, H. *Astr. Astrophys. Suppl.* **70**, 335–367 (1987).
- Parmar, A. N. & White, N. E. in *X-ray Astronomy with EXOSAT* (eds Pallavicini, R. & White, N. E.) (Mem. Astr. Soc. Italia, in the press).
- Paczynski, B. & Sienkiewicz, R. *Astrophys. J.* **268**, 825–831 (1983).
- Rappaport, S., Verbunt, F. & Joss, P. C. *Astrophys. J.* **275**, 713–731 (1983).
- Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296–316 (1979).
- Van den Heuvel, E. P. J., Van Paradijs, J. & Taam, R. E. *Nature* **322**, 153–155 (1986).
- Bhattacharya, D. & Srinivasan, G. *Curr. Sci.* **55**, 327–330 (1986).
- Narayan, R. & Kulkarni, S. *Astrophys. J.* (in the press).
- Cote, J. & Pilyser, E. *Astr. Astrophys.* (in the press).
- Van den Heuvel, E. P. J. *IAU Symp.* **93**, 155–175 (1981).

Interaction between high-velocity pulsar in CTB80 and an infrared emission shell

Robert A. Fesen*, J. Michael Shull*† & Jon M. Saken*

* Center for Astrophysics and Space Astronomy,
University of Colorado, Boulder, Colorado 80309, USA

† Joint Institute for Laboratory Astrophysics, University of Colorado
and National Bureau of Standards, Boulder, Colorado 80309, USA

Possessing an unusual and highly asymmetric radio structure, CTB80 has been classified as a supernova remnant (SNR) with properties resembling those of the Crab Nebula¹⁻³. This interpretation is supported by the presence of a compact X-ray⁴ and radio⁵ source, now known to be a 39.5 ms pulsar^{6,7}, lying near the western edge of CTB80's bright central 'core'. During a survey of infrared emission detected by the Infrared Astronomical Satellite from shock heated dust associated with Galactic SNRs, we discovered a roughly 1° diameter supernova remnant shell centred 30' east of CTB80's core. The pulsar's projected location inside this infrared emission shell, together with similar distance and age estimates, suggest that both the shell and pulsar were produced by the same supernova event. We propose that the interaction between the pulsar's energetic particle emission and the shell's compressed interstellar magnetic field can explain CTB80's remarkable radio structure, and we discuss the implications of such pulsar/SNR interactions.

Using temperature-sensitive 25, 60 and 100 μm ratio maps on the IRAS (Infrared Astronomical Satellite⁸) sky flux images to distinguish shock-heated dust shells associated with known SNRs from other galactic plane sources, we discovered near CTB80 a 64' diameter emission shell centred at $\alpha(1950) = 19\text{ h } 52.9\text{ min}$, $\delta(1950) = 32^\circ 51' (\pm 2')$ [$l = 69.1^\circ$, $b = 2.5^\circ$]. The infrared emission shell can be seen in Fig. 1, which shows a 60 μm image from the sum of three IRAS HCON images, and a 60 μm per 100 μm ratio image constructed by subtracting local background levels on individual 60 μm and 100 μm images, adding the three HCON images together, and applying a linear intensity transformation. As seen in the ratio image, the emission shell appears spherical and nearly complete, except for a break

along its western rim, where there is a faint 20' extension to the southwest.

The shell's radius is $(27.9\text{ pc})D_3$, where D_3 is the shell's distance in units of 3 kpc. Colour-corrected 60 μm and 100 μm fluxes are 600 and 1950 Jy ($\pm 30\%$) respectively, implying a dust temperature of $27.5 \pm 2\text{ K}$ for a λ^{-n} grain infrared emissivity law, where $n = 1.5 \pm 0.5$. Although we could not accurately measure the shell's 12 and 25 μm fluxes, higher resolution IRAS 'co-adds' do indicate some 12 and 25 μm emission. The IRAS fluxes and bandwidths yield a bolometric infrared luminosity of $(5 \times 10^{37}\text{ ergs s}^{-1})D_3^2$.

The detected infrared shell is almost certainly that of an SNR, despite the lack of an identified radio SNR shell in the region⁹. The shell's morphology and dust temperature, together with the presence of faint shock-heated optical filaments^{10,11} coincident with the infrared shell's NE and NW rims and SW extension, all point to a SNR identification. The absence of prominent radio and optical emission, despite a location near the galactic plane, suggests a considerable age for the shell.

We propose that this SNR shell has played an important role in shaping CTB80's remarkable radio morphology. CTB80's radio structure^{1,3,5,9,12,13} consists of three $\sim 30'$ long 'ridges' of nonthermal emission which intersect at a bright $6' \times 10'$ east-west plateau, and along whose western edge lies a smaller but even brighter 'core' of diameter $45''$. CTB80's radio spectral index, which is near zero at the core, steepens to values around -1.0 with increasing distance, with a high degree of polarization. Energetic particle emission from the compact object has been suggested^{12,14,15} as the cause of CTB80's extended radio structure, but the underlying mechanism for its extended three-lobe structure remained unexplained.

A connection between CTB80's extended radio emission and this infrared SNR shell is suggested by the positional coincidences shown in Fig. 2. CTB80's northern arc of nonthermal radio emission matches both the position and curvature of the infrared shell's NW rim, and the bright spot at its northern tip¹⁴ is near a brightening in the shell's infrared emission. Similar correlations can also be seen for the northern ridge's bright region at $\alpha(1950) = 19\text{ h } 50.9\text{ min}$, $\delta(1950) = 32^\circ 51'$, for the sharp boundary just NW of the bright core, and along the SW radio ridge which lies exactly on the shell's faint SW extension. CTB80's pulsar and core are located just inside the shell's western rim with little, if any, coincident infrared emission. The agreement between the reddening values for the shell's NE optical filaments [$E(B - V) \approx 0.8\text{ mag}$]^{2,11} and those in CTB80's optical core,

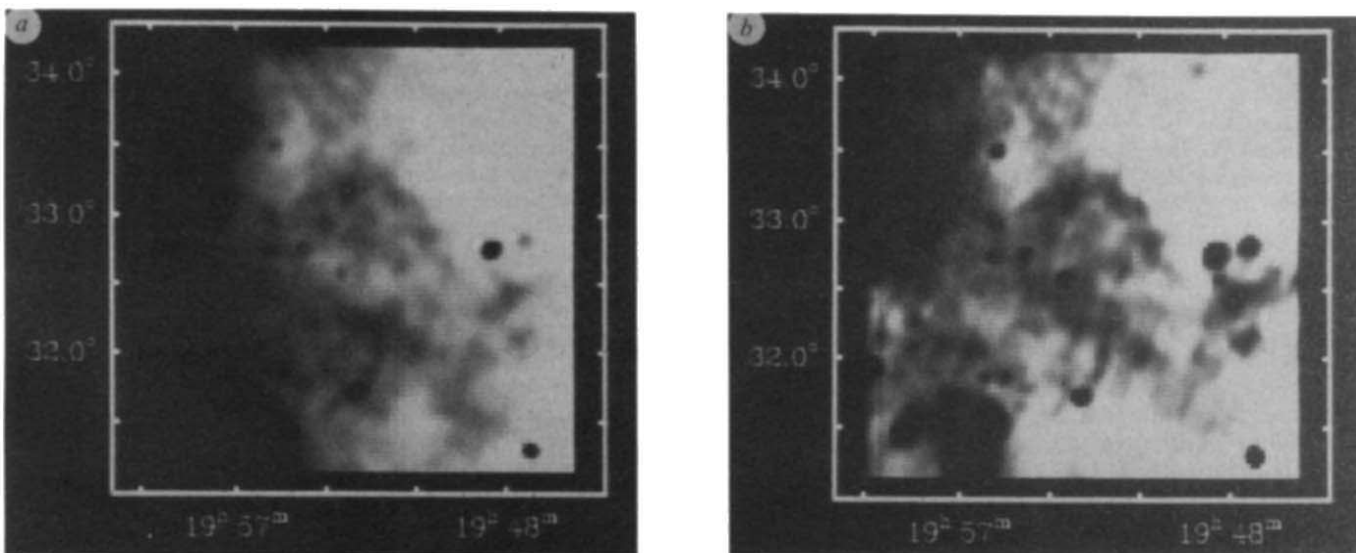


Fig. 1 IRAS Sky Flux HCON sums showing the 60 μm image (a) and a 60 μm per 100 μm flux ratio image (b) for the CTB80 region. The resolution for both images is about 6 arc min.

together with the positional coincidences, also supports their association.

Distance estimates for CTB80 and this infrared shell also appear similar. For a Sedov remnant of age equal to the pulsar's spin-down age⁷ ($P/2\dot{P} = 1.1 \times 10^5$ yr), the shell's distance is $D = (3.5 \text{ kpc})(E_{51}/n_0)^{1/5}t_5^{2/5}$, where E_{51} is the supernova energy in units of 10^{51} ergs, n_0 is the preshock hydrogen density (cm^{-3}), and t_5 is the remnant age in units of 10^5 yr. A distance of (3 kpc) $E_{51}^{0.221} n_0^{-0.257} t_5^{0.3}$ results if one considers the effects of later ('radiative snowplough') evolutionary phases¹⁶ of the SNR. Thus, if E_{51} and n_0 are near unity, the shell's distance is in rough agreement with distance estimates^{2,11} for CTB80 of around 3 kpc. Although the $45 \text{ cm}^{-3} \text{ pc}$ dispersion measure⁷ of pulsar PSR1951+32 implies a distance⁶ of only 1.4 kpc using standard electron density models, the extinction of CTB80's optical core compared with that of neighbouring field stars² suggests a distance in excess of 2 kpc. Distances over 2 kpc and average interstellar hydrogen densities $n_0 \approx 1 \text{ cm}^{-3}$ also produce rough agreement between a swept-up shell mass and the shell's estimated mass from the infrared measurements (see below). Scintillation measurements⁷ on the pulsar imply a transverse velocity of $\sim (300 \text{ km s}^{-1})(D/2 \text{ kpc})^{1/2}$. This velocity puts PSR1951+32 at the high end of the range of pulsar space velocities^{17,18}, and at the extreme limit if $D = 3 \text{ kpc}$. But scintillation-determined velocities are model-dependent and can differ from those determined from proper motions¹⁷.

To explain a physical connection between CTB80 and the SNR shell, as well as the origin of CTB80's peculiar radio emission structure, we propose the following. About 10^5 yr ago, a supernova event occurred which generated both an expanding remnant and a rapidly rotating pulsar. Born with a $\sim 300 \text{ km s}^{-1}$ westward velocity, the pulsar has begun to overtake the decelerating shell. This westward motion is consistent with the radio emission morphology of CTB80's core, the plateau, the eastern ridge and possible optical changes in the core^{3,5,19,20}. The pulsar's space velocity may arise either from a momentum impulse due to its former binary status or from an asymmetry in its relativistic plasma emission²¹. In the former mechanism, the binary must have had a separation $\sim 0.2 \text{ AU}$ and an orbital period ~ 7 days for $V_p \approx 300 \text{ km s}^{-1}$ and a total system mass of $20 M_\odot$. The latter mechanism, although highly speculative, is at least consistent with the close agreement between the pulsar's implied westward motion (position angle $\text{PA} = 85^\circ \pm 3^\circ$) and the pulsar's position along the symmetry axis ($\text{PA} = 86^\circ \pm 3^\circ$) of the core's Balmer-dominated east and west optical emission lobes^{22,23}. These emission lobes exhibit a bow-shock morphology suggestive of the working surfaces of jets.

The time needed for a 300 km s^{-1} pulsar to traverse the shell's radius is $(9 \times 10^4 \text{ yr})D_3$. This is remarkably close to the pulsar's spin-down age. The pulsar's predicted angular proper motion is $(0.021 \text{ arc s yr}^{-1})D_3^{-1}$ in a westward direction. We note that the required position angle ($\text{PA} = 85^\circ \pm 3^\circ$) to move the pulsar back through the shell's centre is exhibited by the sharp southern boundary of CTB80's eastern radio ridge. This qualitatively supports the suggestion⁶ for a trail of relativistic particles as a result of the pulsar's westward motion, but CTB80's bright plateau immediately adjacent to the pulsar and core does not show this alignment.

The pulsar's position just inside the shell's western edge suggests that CTB80's extended N and SW radio emission ridges result from the injection of pulsar-generated relativistic particles onto the shell's compressed magnetic field lines. Consequently, the resulting synchrotron emission follows the shell's circular geometry to the north and its breakout to the SW. The spectral index variation away from the pulsar would then result from the expected particle energy losses. We believe it unlikely that the observed nonthermal radio flux in these ridges is conventional SNR radio emission, because of the symmetrical position of the NE and SW ridges relative to the pulsar, as well as the systematic spectral index variations. Although shock-heated

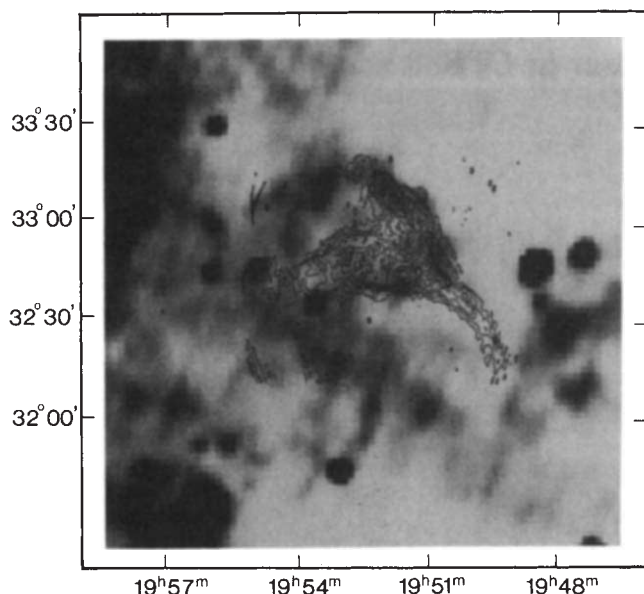


Fig. 2 CTB80's nonthermal radio emission contours³ superimposed on the $60 \mu\text{m}$ per $100 \mu\text{m}$ ratio image, showing the positional coincidence of the NW and SW radio ridges with the infrared emission shell, as well as the location of the core just inside inner western rim of the shell. The positions of the NE shock-heated optical filaments^{10,11} are also indicated.

optical filaments along these ridges indicate the likelihood of some shell-generated radio emission, its contribution is probably small. Little, if any, radio emission is detected at the shell's brightest optical filaments along the northeastern rim⁹. In this view, CTB80's fairly straight eastern radio ridge results from the pulsar's relativistic particle wake inside the SNR's swept-out cavity. This is consistent with the lack of associated shock-heated optical emission in the eastern radio ridge^{10,11}.

Although such a large SNR is likely to have entered the 'radiative snowplough' phase, we can gain some insight by considering the adiabatic Sedov model, in which a consistent picture emerges for the remnant's size, preshock density and distance. The shell velocity is then given by $V_s = (123 \text{ km s}^{-1})(E_{51}/n_0)^{1/5}t_5^{-3/5}$ and the post-shock temperature by $T_s = (2 \times 10^5 \text{ K})(E_{51}/n_0)^{2/5}t_5^{-6/5}$. A shock velocity $\sim 85 \text{ km s}^{-1}$ is suggested by a comparison of the shell's optical filament spectra¹¹ to shock emission models²⁴. (Note: $V_s = 80 \text{ km s}^{-1}$ is also consistent with the radiative snowplough solution at $t_5 = 1$; however, the optical filaments may be excited by slower shocks.) The $P-V$ work done on the shell is $(5 \times 10^{38} \text{ erg s}^{-1})E_{51}t_5^{-1}$, which, when compared to the shell's infrared luminosity, indicates a conversion efficiency of mechanical energy into infrared dust emission of $\sim 10\%$. Using standard infrared dust emissivities²⁵, we estimate a mean hydrogen density $n_H \approx 3 \text{ cm}^{-3}$ in the shell and a mass of dust $M_d \approx (6 M_\odot)D_3^2$. If 50% of the grains are destroyed by sputtering and if the dust-to-gas ratio²⁵ is 0.0075, the shell mass $M_{\text{gas}} \approx (1,600 M_\odot)D_3^2$, which is comparable to the $(4,450 M_\odot)E_{51}^{3/5}n_0^{2/5}t_5^{6/5}$ of swept-up interstellar matter estimated from the Sedov model. Even better agreement would result if E_{51} , n_0 , or t_5 were slightly less than 1. Alternatively, the dust may be partially heated by interstellar radiation, in which case, dust temperatures and masses will be altered.

In our picture, CTB80's SNR classification is not that of a centrally peaked or 'plerion' SNR but rather a 'composite-type', as it consists of both an outer expanding shell and an inner flat-spectrum radio emission region surrounding its associated pulsar. It differs from other composite SNRs only in that its

central source is situated close to the remnant's expanding outer shell and is currently interacting with that shell. Despite having a rapid pulsar, CTB80's probable age of $\sim 10^5$ yr makes it quite different from the Crab Nebula, and also unlikely to be related to a possible AD 1408 historical SN, as has been proposed elsewhere²⁶.

A similar, but more evolved, version of CTB80 may be the radio source G5.4-1.2. This object exhibits a shell-like structure with a small, flat-spectral-index source (G5.27-0.90) just outside, but connected to its western arc of emission by a small narrow ridge of emission^{27,28}. The recent discovery²⁹ of a 125 ms pulsar near G5.27-0.90, together with the radio emission's symmetry line with respect to this source, is reminiscent of CTB80's interaction between energetic particles and the remnant shell's compressed magnetic field lines. In this case, however, the pulsar may have penetrated through the SNR shell.

Such pulsar-shell interactions should be infrequent, owing to the small percentage of pulsars with sufficient velocities to catch-up to their expanding shell and to the short encounter lifetimes. The time for a pulsar with velocity $V_p = (100 \text{ km s}^{-1}) V_{100}$ to catch up with a Sedov-type shell is $t_c = (6.5 \times 10^5 \text{ yr}) (E_{51}/n_0)^{1/3} V_{100}^{-5/3}$. If a SNR shell remains intact for $\sim 2 \times 10^5$ yr, then only pulsars with velocities in excess of about $(200 \text{ km s}^{-1}) (E_{51}/n_0)^{1/5}$ can catch up with their related SNR shells. From scintillation¹⁷ and proper motion¹⁸ surveys, only $\sim 10\%$ of the measured pulsars have velocities greater than 200 km s^{-1} .

The lifetime for pulsar-shell interaction is equally restrictive. If a pulsar must lie within a distance $\Delta r = (1 \text{ pc}) r_{pc}$ on either side of the shell, in order for the pulsar's energetic particles to interact with the shell's magnetic field, then the interaction lifetime is $\Delta t = (2 \times 10^4 \text{ yr}) r_{pc} V_{100}^{-1}$. Assuming a pulsar B-field decay time of 2×10^6 yr and that only a fraction, f_{fast} , of all pulsars are fast enough to catch an intact SNR shell, then the fraction of pulsars capable of such encounters is $10^{-2} r_{pc} V_{100}^{-1} f_{fast}$. If we set $V_{100} \approx 2$ and $f_{fast} \approx 0.1$ as discussed above, then the possible existence of two interactions (CTB80 and G5.4-1.2) among the approximately 500 known galactic pulsars implies that $r_{pc} \approx 8 \text{ pc}$. Alternatively, r_{pc} could be smaller if some SNR shells remain intact for 5×10^5 yr or if the fraction of fast pulsars exceeds 10% .

We thank R. Becker, E. Dwek, A. Hamilton, J. Hester and C. Kennel for useful discussions. This research was partially funded by IRAS and Astrophysical Theory grants from NASA.

Received 29 February; accepted 5 May 1988.

- Velusamy, T. & Kundu, M. R. *Astr. Astrophys.* **32**, 375-390 (1974).
- Angerhofer, P. E., Wilson, A. S. & Mould, J. R. *Astrophys. J.* **236**, 143-152 (1980).
- Angerhofer, P. E., Strom, R. G., Velusamy, T. & Kundu, M. R. *Astr. Astrophys.* **94**, 313-322 (1981).
- Becker, R. H., Helfand, D. J. & Szymkowiak, A. E. *Astrophys. J.* **255**, 557-563 (1982).
- Strom, R. G. *Astrophys. J. Lett.* **319**, L103-L108 (1987).
- Kulkarni, S. R. *et al. Nature* **331**, 50-53 (1988).
- Fruchter, A. S. *et al. Nature* **331**, 53-54 (1988).
- Neugebauer, G. *et al. Astrophys. J. Lett.* **278**, L1-L6 (1984); **217**, 425-433 (1977).
- Mantovani, F., Reich, W., Salter, C. J. & Tomasi, P. *Astr. Astrophys.* **145**, 50-58 (1985).
- van den Bergh, S. *Publ. astr. Soc. Pacif.* **92**, 768-770 (1980).
- Blair, W. P., Kirshner, R. P., Fesen, R. A. & Gull, T. R. *Astrophys. J.* **282**, 161-171 (1984).
- Sofue, Y., Takahara, F., Hirabayashi, H., Inoue, M. & Nakai, N. *Publ. astr. Soc. Japan* **35**, 437-445 (1983).
- Strom, R. G., Angerhofer, P. E. & Dickel, J. R. *Astr. Astrophys.* **139**, 43-49 (1984).
- Velusamy, T. & Kundu, M. R. *J. Astrophys. Astr.* **4**, 253-260 (1983).
- Wang, Z. R. & Seward, F. D. *Astrophys. J.* **285**, 607-612 (1984).
- Cioffi, D. F., McKee, C. F. & Bertschinger, E. *Astrophys. J.* **334** (in the press).
- Cordes, J. M. *Astrophys. J.* **311**, 183-196 (1986).
- Lyne, A. G., Anderson, B. & Salter, M. J. *Mon. Not. R. astr. Soc.* **201**, 503-520 (1982).
- Dickel, J. R., Angerhofer, P. E., Strom, R. G. & Smith, M. D. *Vistas Astr.* **25**, 127-135 (1981).
- Strom, R. G. & Blair, W. P. *Astrophys. J.* **149**, 259-265 (1985).
- Harrison, E. & Tademaru, E. *Astrophys. J.* **201**, 447-461 (1975).
- Fesen, R. A. & Gull, T. R. *Astrophys. Lett.* **24**, 197-204 (1985).
- Hester, J. J. & Kulkarni, S. R. *Bull. Amer. astr. Soc.* **19**, 1088 (1987).
- Cox, D. P. & Raymond, J. C. *Astrophys. J.* **298**, 651-659 (1985).
- Dwek, E. *Astrophys. J.* **322**, 812-821 (1987).
- Strom, R. G., Angerhofer, P. E. & Velusamy, T. *Nature* **284**, 38-40 (1980).
- Becker, R. H. & Helfand, D. J. *Nature* **313**, 115-118 (1985).
- Caswell, J. L. *et al. Mon. Not. R. astr. Soc.* **225**, 329-334 (1987).
- Manchester, R. N., D'Amico, N. & Tuohy, I. R. *Mon. Not. R. astr. Soc.* **212**, 975-986 (1985).

The parent structure of the layered high-temperature superconductors

T. Siegrist, S. M. Zahurak, D. W. Murphy & R. S. Roth*

AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

* National Bureau of Standards, Gaithersburg, Maryland 20899, USA

Oxide superconductors in the system $\text{Ti}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ (ref. 1) have transition temperatures (T_c) above 100 K, increasing with n . So far, stacking sequences up to $n = 3$ have been found in small crystals, and sequences with $n > 3$ have been seen in electron microscopy studies¹. For large n , the stoichiometry of $\text{Ti}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ approaches CaCuO_2 , a structure expected to consist only of CuO_2 planes separated by Ca atoms. By analogy with $\text{Ti}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$, the unit cell of this hypothetical phase is expected to be tetragonal with $a = 3.86 \text{ \AA}$. Such a compound is not known in the Ca-Cu-O system, but Roth² recently reported that small amounts of Sr on the Ca site can stabilize this simple structure. Here we report the growth of small single crystals of this phase, with composition $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$, and their characterization by single-crystal X-ray diffraction. The crystals are tetragonal with space group $P4/mmm$, and the structure contains planar $[\text{CuO}_2]_\infty$ layers separated by Ca and Sr atoms. The structure is a simple defect perovskite with ordered oxygen vacancies and can be regarded as the $n = \infty$ parent of the $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ ($\text{A} = \text{Bi, Tl; B} = \text{Sr, Ba}$) superconductors.

Crystals were prepared from stoichiometric amounts of $\text{Ca}(\text{OH})_2$, $\text{Sr}(\text{NO}_3)_2$, and CuO (0.85 Ca: 0.15 Sr: Cu) heated to 1,423 K. Melting occurs and the growth of small platelets is promoted. The platelets show a metallic lustre, are brittle and usually only a few microns thick. They were examined on an ENRAF-NONIUS CAD4 single-crystal diffractometer. All reflections in an Ewald sphere up to $2\theta = 80^\circ$ were measured. All subsequent calculations were carried out using the NRCVAX structure package on a MicroVax I computer³. Details of the analysis are given in Table 1, atomic positions in Table 2 and temperature parameters in Table 3. The lattice parameters were determined by measuring absolute 2θ values of 32 reflections in the range of 50° to 65° . To address the ambiguity in the space group, refinements in $P4m2$ (non-centric) and $P4/mmm$ (centric) were carried out. The space group $P4m2$ allows the oxygen

Table 1 Summary of crystallographic information for $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$

Dimensions (mm)	0.06 × 0.05 × 0.004
Radiation	MoK α
Monochromator	Graphite
Crystal system	Tetragonal
Space group	$P4/mmm$ (No. 123)
Temperature	Ambient
Calculated density	4.95 g cm ⁻³
Scan mode	$2\theta - \omega$
2θ range	0 to 80°
μ (cm ⁻¹)	170.7
Absorption correction	Gaussain integration
Transmission factors	0.53 to 0.94
Extinction (μm)	4.1 (2)
Total reflections	1,516
Unique reflection	112
Observed reflection ($I > 2\sigma$)	106
Agreement factor on averaged intensities	0.03
Data/parameter	10.6
R_f	0.035
R_w	0.040

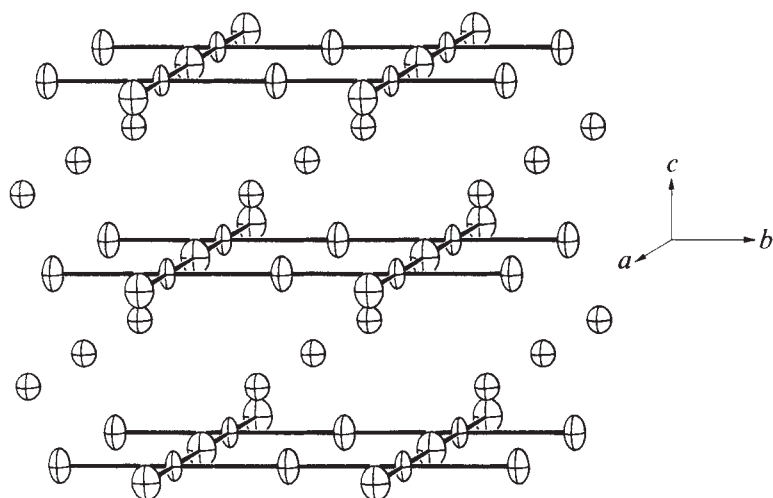


Fig. 1 Extended view of $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$. Only the Cu-O bonds are drawn.

atom to move in the z -direction, but refinements in this space group did not produce significant deviations from $z=0$. Two local minima, slightly above and below $z=0$, were found in the refinement in $P4m2$. Disorder of the oxygen atoms due to a local distortion produced by the difference in size between Ca and Sr may well account for this feature. As a result of the small scattering power of oxygen, the refined position parameter obtained is barely reliable. Further investigations using neutron powder diffraction, where the oxygen scattering power is enhanced, are planned. For any further refinement based on X-ray data, we chose the space group $P4/mmm$ for the final refinement, where the oxygen position is fixed in $(0, \frac{1}{2}, 0)$ and an anisotropic temperature parameter can be refined.

The structure, as shown in Fig. 1, is very simple. The $[\text{CuO}_2]_\infty$ planes are flat, the puckering that occurs for many of the layered copper oxides was not observed. But flat $[\text{CuO}_2]_\infty$ layers occur in the $\text{Tl}-\text{Nd}_2\text{CuO}_4$ -type structure⁴ and in $\text{Ti}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ with $n=1$ and $n=3$ (ref. 1). For $n=3$, only the middle $[\text{CuO}_2]_\infty$ layer is flat, whereas the adjacent layers are slightly puckered. The structure also has the same cation skeleton as $\text{Ba}_2\text{YCu}_3\text{O}_6$, but differs in the ordering of oxygen. Pyramidal coordinated copper and linearly two-coordinated copper is found in $\text{Ba}_2\text{YCu}_3\text{O}_6$ (ref. 5). Moving the apical oxygen in $\text{Ba}_2\text{YCu}_3\text{O}_6$ from $(0, 0, z)$ to $(0, \frac{1}{2}, 0)$ produces $[\text{CuO}_2]_\infty$ layers. Although both structures have the empirical formula ACuO_2 , they differ in the formal charge of Cu. The copper is formally 2^+ in $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$ and 1.67^+ in $\text{Ba}_2\text{YCu}_3\text{O}_6$. The driving force

for the different oxygen ordering is to create distinct Cu^{2+} (square pyramidal) and Cu^{1+} (linear-coordinate) sites in $\text{Ba}_2\text{YCu}_3\text{O}_6$, whereas all the Cu sites are equivalent and suitable for Cu^{2+} in $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$.

In $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$, the Cu-O distance of 1.931 Å compares well with the distances found in other cuprates with planar $[\text{CuO}_2]_\infty$ sheets^{1,4-7}. The $[\text{CuO}_2]_\infty$ sandwich Ca (Sr) in a way similar to that in $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ and $\text{Ba}_2\text{YCu}_3\text{O}_7$, with a Ca(Sr)-O distance of 2.507 Å. The Ca(Sr) coordination polyhedron is a square prism (coordination number 8) with a length-to-base ratio of 1.172, approaching a cube. The distance between the $[\text{CuO}_2]_\infty$ sheets is slightly shorter than in $\text{Ba}_2\text{RCu}_3\text{O}_7$ (ref. 7) and $\text{Ti}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ (ref. 1), but is expected to vary with the Ca-to-Sr ratio. In addition, mismatch between the Ti_2O_2 layers and the $[\text{CuO}_2]_\infty$ layers in $\text{Ti}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ may be accommodated by puckering of the $[\text{CuO}_2]_\infty$ sheets.

This structure can be regarded as the parent structure of $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ for large n in the case of the $\text{Tl}-\text{Ba}-\text{Ca}-\text{Cu}-\text{O}$ (ref. 1) and $\text{Bi}-\text{Sr}-\text{Ca}-\text{Cu}-\text{O}$ (refs 8-11) superconductor systems. The absence of Ti_2O_2 (Bi_2O_2) layers results in a regular coordination for Ca (Sr) and Cu, and gives rise to a high symmetry and a small unit cell volume. As only the $[\text{CuO}_2]_\infty$ layers are present in this new phase, their properties can now be studied in detail without the interference of the Ti_2O_2 (Bi_2O_2) layers in $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$, or the CuO_3 chains in $\text{Ba}_2\text{YCu}_3\text{O}_7$. The valence state of copper in this phase is 2^+ and therefore the compound is an insulator, as confirmed by resistivity measurements on powder samples. Attempts to introduce charge carriers by doping that could render the compound metallic and lead to superconductivity are being pursued.

In summary, we have characterized the structure of $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$, a compound built from CuO_2 sheets that sandwich Ca(Sr). This arrangement is the parent structure of the superconducting system $\text{A}_2\text{B}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{4+2n}$ with $\text{A}=\text{Tl}$, Bi , $\text{B}=\text{Ba}$, Sr for large n , where additional CaCuO_2 layers are added.

Received 13 June; accepted 27 June 1988.

1. Torardi, C. C. *et al. Science* **240**, 632-633 (1988).
2. Roth, R. S. *Am. Phys. Soc. Meeting, New Orleans* (1988).
3. Gabe, E. J., Lee, F. L. & Le Page, Y. in *Crystallographic Computing* Vol. 3 (ed. Sheldrick, G. H., Krüger, C. & Goddard, R.) 167-174 (Clarendon, Oxford, 1985).
4. Müller-Buschbaum, H.-K. & Wollschläger, W. *Z. Anorg. Allg. Chem.* **414**, 76-80 (1975).
5. Gallagher, P. K., O'Brian, H. M., Sunshine, S. A. & Murphy, D. W. *Mater. Res. Bull.* **22**, 995-1006 (1987).
6. Müller-Buschbaum, H.-K. *Angew. Chem.* **89**, 704-717 (1977).
7. Le Page, Y. *et al. Phys. Rev. B* **36**, 3617-3621 (1987).
8. Sunshine, S. A. *et al. Phys. Rev. B* (in the press).
9. Von Schnering, H. G. *et al. Angew. Chem. Int. Ed. Engl.* **27**, 574-576 (1988).
10. Tarascon, J. M. *et al. Phys. Rev. B* (in the press).
11. Subramanian, M. A. *et al. Science* **239**, 1015-1017 (1988).

Table 2 Crystallographic data of $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$ tetragonal cell, space group $P4/mmm$ (No. 123), $Z=1$, $a=3.8611(2)$ Å, $c=3.1995(2)$ Å

Atom	Position	x	y	z	B_{iso} (Å ²)	Occ
Ca	1d	1/2	1/2	1/2	0.52 (4)	0.86 (2)
Sr	1d	1/2	1/2	1/2	0.52 (4)	0.14 (2)
Cu	1a	0	0	0	0.40 (3)	
O	2f	0	1/2	0	0.69 (17)	

Table 3 Anisotropic thermal parameters of $(\text{Ca}_{0.86}\text{Sr}_{0.14})\text{CuO}_2$ $B = -2\pi^2(h^2u_{11}a^{*2} + \dots)$; $u_{ij} = 0$ for $i \neq j$

Atom	u_{11}	u_{22}	u_{33}
Ca (Sr)	0.62 (5)	0.62	0.75 (7)
Cu	0.27 (3)	0.27	0.97 (5)
O	0.88 (23)	0.46 (20)	1.30 (22)

Effect of climate change on fire regimes in northwestern Minnesota

James S. Clark*

Department of Ecology, 318 Church Street SE,
University of Minnesota, Minneapolis, Minnesota 55445, USA

Of all the impacts of projected climate change on forest ecosystems, perhaps the most difficult to forecast is the potential for altered fire frequency and intensity. Fire regimes in forests are poorly understood for lack of long-term evidence. Here I used petrographic thin sections to determine the annual production of charcoal within a lake catchment in northwestern Minnesota over the past 750 years providing the long and high-resolution record required to elucidate fire regimes. Maximum abundance and frequency occurred in the warm, dry fifteenth and sixteenth centuries. Fire importance decreased dramatically with the onset or intensification of the 'little ice age' about AD 1600. Fire cycles with harmonics corresponding to multiples of the 22-year drought cycles of the region and increased fire frequency at times when early successional stands were breaking up, suggest a synergistic influence of climate and fuel accumulation. The anomalously warm, dry twentieth-century climate would have produced substantially different fire regimes from the previous century in the absence of fire suppression.

The study area near the prairie-forest border in Itasca State Park supports old-growth red pine (*Pinus resinosa*) trees with fire scars, indicating frequent fire before the advent of fire suppression in AD 1910 (ref. 1). Trees of aspen (*Populus tremuloides*, *P. grandidentata*) and paper birch (*Betula papyrifera*) that are 80–90 years old comprise much of the remaining canopy. Available data from the northern Great Plains² suggest relatively warm, dry fifteenth and sixteenth centuries, separating cooler and wetter times from the onset of the 'little ice age' that followed³ (Fig. 1a). On a finer scale is the 22-year drought cycle, known from documents in historic times⁴ and producing a significant periodicity in tree-ring drought indices since 1700 (the length of the record)⁵.

Petrographic thin sections prepared from cores of annually laminated lake sediments permit estimates of annual charcoal delivery and avoid the problems associated with other methods of stratigraphic charcoal analysis⁶. The method involves dehydration of sediment samples in acetone, impregnation with epoxy resin, and sectioning by petrographic methods⁷. Because only relatively large fragments are counted, charcoal diagrams reflect local fire⁸. Varve counts are believed reliable to ± 5 years; varves crossdated among three lakes for the period 1640–1800 (J.S.C., manuscript in preparation).

Charcoal data indicate that fire regimes responded to climatic changes of the past few centuries. Charcoal was most abundant at Deming Lake when warmer, drier conditions prevailed from AD 1400 to 1600 (Fig. 1b). The little ice age saw a decreased importance of fire, as witnessed both by lower mean charcoal amount ($p < 0.01$, *K-S* 2-sample test) and lower frequency of large charcoal peaks; a periodicity is represented by maxima near AD 1620, 1705, 1790 and 1880. Charcoal analyses of sediments dating from 1640 to 1985 from two nearby lakes display a similar pattern, and fire-scar data spanning the period since 1650 substantiate the hypothesis that maxima represent times of increased fire⁸. The almost complete absence of charcoal accompanying the fire-suppression period begun in 1910 indicates that charcoal data faithfully depict short-term changes in fire regimes. Fourier analysis of the detrended (with a cubic polynomial) charcoal series indicate 80–95-year ($p < 0.05$), 33–

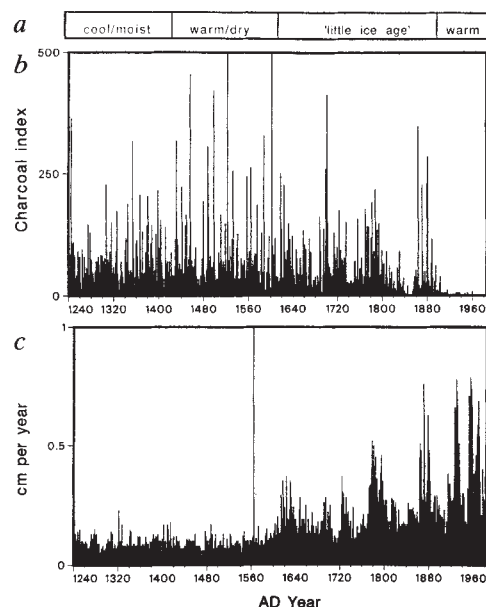


Fig. 1 General climatic conditions *a*, for the northern Great Plains since AD 1240, together with charcoal accumulation, *b*, and varve thicknesses, *c*, for the same period from Deming Lake. The charcoal index is $(\text{cm}^2 \text{ charcoal} / \text{cm}^2 \text{ sediment year}) \times 100$, determined from thin sections.

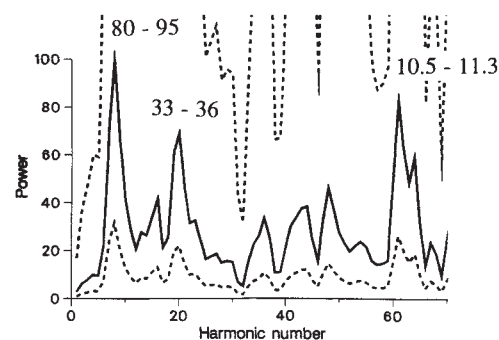


Fig. 2 Fourier series analysis of charcoal series in Fig. 1b with 95% confidence intervals (dashed lines). Years corresponding to significant ($p < 0.5$) harmonics are indicated.

36-year ($p < 0.1$), and 11-year ($p < 0.01$) harmonics for the entire series (Fig. 2). A 43–46-year harmonic is also highly significant ($p < 0.01$) for the period since 1400, when warm, dry conditions began. Recent large charcoal maxima are in phase with several drought periods in the 1890s, 1800 and 1730 (ref. 5).

An explanation for these cycles derives from the two fundamental processes affecting fire frequency: fuels and weather. For roughly 30 to 60 years following fire, the fuel loading is generally insufficient to support more than light surface fires^{9–11}. Although fire may produce more fuel than it consumes, these fuels are highly discontinuous spatially. Eventually a more continuous litter layer, increased production of coarse woody debris, and increased vertical continuity of fine fuels contributed by regrowth of vegetation cause an increase in the probability of intense fire. Fuel continuity and ability to support a fire increase in a logistic fashion and, depending on the region, achieve some maximum level 150 to 300 years following the last fire^{12–14}. Olsen¹⁵ suggests an equivalent numerical model for fire probability as a function of time since the last fire. Weather is highly stochastic, but the well documented drought cycle in the Great

* Present address: The Biological Survey, New York State Museum, Albany, New York 12230, USA.

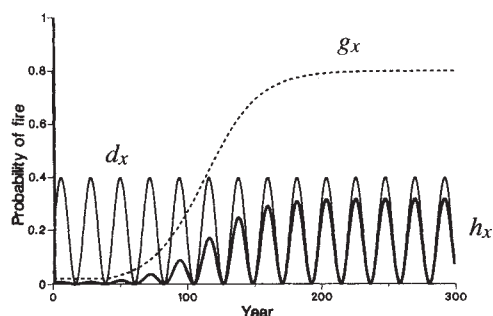


Fig. 3 Model of fire probability as a function of time since the last fire. The function describing the fuel contribution to probability of fire, $g_{x+1} = g_x \exp[\gamma(1 - g_x\delta)]$, where γ is a rate constant and δ is the maximum probability obtained at high fuel loadings, reflects discontiguous loadings for a period of several decades at which point increased contributions of coarse woody debris increase, multiple fuel layers develop, and forest-floor fuels become more continuous. The heavier fuels and understory shrubs represent a substantial fuel hazard during extended dry periods that follow drought cycles in the region $d_x = \alpha \sin[2\pi x/\lambda] + E[d_x]$, with amplitude α , wavelength $\lambda = 22$ yr, and expectation $E[d_x]$, being less responsive to short-term weather fluctuations. The composite fire probability function, h_x , concentrates high-probability fire years at higher multiples of the 22-yr drought cycle, thus providing a mechanism for low-frequency fire cycles observed in the charcoal series.

Plains and adjoining states⁵ represents a substantial component of weather than can be treated deterministically. By affecting the proportion of existing organic fuels that can burn¹⁰⁻¹¹, drought years exercise an important control on fire frequency and intensity, particularly depth of burn.

But neither of these processes alone results in the observed cycles—the drought cycle predicts fire years at 22-year intervals, and the fuel model results in a generally increasing probability of fire with time, but with high variance. Two factors, drought cycles and break-up of early successional species, could serve to cluster high-probability fire years at predictable intervals corresponding to the observed frequencies. The probability of fire in any given year is the product of the two marginal densities, incorporating processes of fuel accumulation and changing probability of fire determined by drought cycles. The fact that neither process alone produces cycles less than 22 years is readily demonstrated by reducing the importance of the drought cycle (decreasing either the amplitude or expected mean probability), which removes the periodicity. Obviously, without the fuel function there remains only a 22-year cycle. This model predicts that high-probability fire years should be clustered at 33–44, 55–66 and 77–88 years (Fig. 3), namely intervals of increasing moisture stress. The importance of this mechanism and the realized interval would depend on interaction of the drought cycle and the shape of the fuel function.

Pollen diagrams from this and two adjacent small lakes with varved sediments indicate that high-charcoal intervals since AD 1640 (1770–1810, 1870–1890) coincided with times when aspen and paper birch stands were breaking up (J.S.C., manuscript in preparation), suggesting a second synchronizing factor. This would provide a concentration of fuels at 80–90-year intervals that could burn when weather conditions were suitable¹²⁻¹⁴. These species were recruited following intense fires of the previous cycle. The existing overmature stands of these species on the study area might have represented the most recent cycle of burning if fire suppression were not in practice. Thus patterns of tree recruitment following fire could serve to reinforce fire cycles. The observed periodicities might result from the fact that fires are most likely to occur at multiples of the 22-year drought cycle, and that fires are more severe at higher multiples of the cycle as a consequence of fuel build-up.

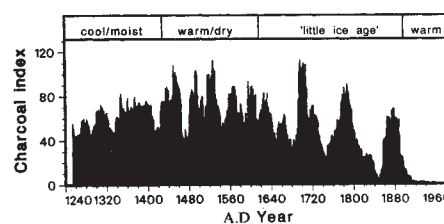


Fig. 4 Charcoal series presented in Fig. 1b filtered with a 15-yr running mean. The 44-yr cycles of the warm, dry fifteenth and sixteenth centuries shift to 88-yr cycles with the cooler, moister conditions that set in about 1600.

The model suggests the potential influence of changing climate on fire regimes. By modifying the initial lag period before accumulating fuels and vegetation structure result in a rapid increase in fire probability, and by influencing production rates and thus the rate at which fire probabilities increase, a changing climate could result in fire cycles that shift among multiples of the 22-year drought cycle. This prediction is supported by the original data. A filtered series (Fig. 4) indicates a shift from a 44-year cycle that prevailed during the warm, dry fifteenth and sixteenth centuries to an 80–95-year cycle with the onset of the little ice age. The lower decomposition rates that would have characterized cooler conditions¹⁶ could permit a build-up of moist humus that would not burn in surface fires, which consume litter at shorter intervals¹⁰. More intense fires would be possible during dry years, thus emphasizing longer cycles.

Analysis of annual charcoal accumulation in varved lake sediments reveals fire cycles and climatically induced changes in fire regimes that were previously unknown. Cycles have been proposed but not demonstrated. There has been no serious consideration of recent climatic change as a factor in fire regimes. Fire cycles demonstrated here appear to be driven by climate and the dynamics of fuel regimes, consisting of short-term, less intense fires superimposed on less frequent intense burns. The scale at which these changes have taken place is a matter of concern, for changing fire regimes over decades to centuries impact forest trees now on the landscape. Many extant forests regenerated when fire regimes were very different (such as during the little ice age) from those prevailing at the times of European settlement or the advent of fire suppression. Efforts to manage forests and National Parks consistent with the presettlement condition must recognize that these short-term climatic events can have profound influence on fire regimes. The 'natural' regime depends on the point in the past chosen for emulation. These results suggest that it could be difficult or impossible to resurrect earlier fire regimes in the twentieth century with its anomalous climate, because these climatic changes in the past have meant different fire regimes. Indeed, increased evapotranspiration and decreased precipitation of the order observed in this century would probably have resulted in at least a 25% increase in fire frequency in the absence of fire suppression (J.S.C., manuscript in preparation). With continued fire suppression and the further warming, fuel build-up will result in more intense and/or more frequent fire.

I thank P. Abrams, C. Bingham, E. J. Cushing, D. Grigal, G. King, R. Skaggs, D. Tilman and H. E. Wright for discussion. Confidence-interval estimates for Fourier series were suggested by C. Bingham. Financial support was provided by the Dayton-Wilkie Foundation, the National Science Foundation, the Whitehall Foundation, the Academic Computing Center of the University of Minnesota, and a Doctoral Dissertation Fellowship from the University of Minnesota. Contribution numbers for this work were 351 of the Limnological Research Center and 574 of the New York State Science Service.

Received 23 February; accepted 31 May 1988.

1. Frissell, S. S. *Quat. Res.* **3**, 397–407 (1973).
2. Leonard, E. M. *Quat. Res.* **25**, 199–214 (1986).
3. E. C. Grimm *New Phytol.* **93**, 311–350 (1983).
4. Borchert, J. R. A. *Ann. Am. Geogr.* **61**, 1–22 (1971).
5. Mitchell, J. M., Stockton, C. W. & Meko, D. M. in *Solar-Terrestrial Influences on Weather and Climate* (eds McCormac, B. M. & Seliga, T. A.) 125–143 (Reidel, Dordrecht, 1979).
6. Clark, J. S. *Quat. Res.* **30**, 81–91 (1988).
7. Clark, J. S. *Quat. Res.* **30**, 67–80 (1988).
8. Clark, J. S. *Ecol. Monogr.* (in the press).
9. Heinselman, M. L. *Quat. Res.* **3**, 329–382 (1973).
10. Haines, D. A., Main, W. A. & McNamara, E. F. *USDA For. Serv. Res. Pap. NC-158* (1978).
11. Chandler, C., Cheney, P., Thomas, P., Trabaud, L. & Williams, D. *Fire in Forestry I: Forest Fire Behavior and Effects* (Wiley, New York, 1983).
12. Gore, J. A. & Patterson, W. A. *Can. J. For. Res.* **16**, 335–339 (1986).
13. Harmon, M. E. *et al. Adv. Ecol. Res.* **15**, 133–302 (1986).
14. Heinselman, M. L. in *Fire Regimes and Ecosystem Properties* (eds Mooney, H. A., Bonnicksen, T. M., Christensen, N. L., Lotan, J. E. & Reiners, W. A.) 7–57 (USDA For. Serv. Gen. Tech. Rep. GTR-WO-26, Washington DC, 1981).
15. Olsen, J. S. in *Fire Regimes and Ecosystem Properties* (eds Mooney, H. A., Bonnicksen, T. M., Christensen, N. L., Lotan, J. E. & Reiners, W. A.) 327–378 (USDA For. Serv. Gen. Tech. Rep. GTR-WO-26, Washington DC, 1981).
16. Meentemeyer, V. *Ecology* **59**, 465–472 (1978).

Collapse of a Hercynian Tibetan Plateau into a late Palaeozoic European Basin and Range province

Gilles Ménard* & Peter Molnar*†

* Laboratoire de Géophysique Interne et Tectonophysique, Institut de Recherches Interdisciplinaires de Géologie et de Mécanique, Université Joseph Fourier, BP 53X, 38041 Grenoble Cedex, France

† Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The large width, the importance of crustal shortening and the likelihood of thick crust in the late Palaeozoic Hercynian (or Variscan) belt of southern Europe suggest a tectonic evolution analogous to the evolution of the Tibetan Plateau and its neighbouring Himalaya¹, and of the Altiplano and other high plateaux of the Andes. Here we extend Dewey and Burke's analogy¹ of the Hercynian chain with Tibet to include the subsequent stage of pervasive normal faulting and crustal extension and thinning, which characterize the active tectonics of Tibet, parts of the high Andes, and the Basin and Range province of western North America. It appears that normal faulting and rifting also characterize the late Palaeozoic tectonic activity of the Hercynian chain.

The Hercynian belt of southern Europe comprises a roughly easterly-trending zone from the Iberian Peninsula (when reconstructed to France by closing the Bay of Biscay) across most of France and southern Britain, and through Germany and the basement of the Alps into eastern Europe. This belt is commonly associated with widespread, roughly north-south, late Palaeozoic crustal shortening, which probably occurred during a collision of Gondwana with Laurasia during the middle to late Palaeozoic era^{2–5}. The involvement of the basement in many of the thrust sheets in the Hercynian belt⁶ requires that the deformation include the entire crust of a broad area, and not just a detached sedimentary cover. The large magnitudes of thrust displacement on several thrust faults^{2–6} imply that the crustal thickness of the Hercynian belt increased during this stage of deformation. The breadth of the Hercynian belt, the thickness and sedimentary facies of its cover, and the high metamorphic grade^{3,5} of rock involved in the thrust faulting imply that initially the crust was not unusually thin. Thus the crust of the Hercynian chain was almost surely thickened significantly during crustal shortening⁵.

Whether the Hercynian crustal thickness reached 60–70 km, as it does beneath present-day Tibet⁷ and the Andes⁸, cannot

be determined. Conversely, whereas the evidence for abundant crustal shortening in the Hercynian chain is clear and has been likened to that of the Himalaya^{3,5}, the evidence for important crustal shortening in Tibet⁹ or the Andes¹⁰ is less widely accepted.

The Hercynian chain is renowned for widespread granites, many of which were intruded late in the development of the chain^{5,11}. By no means can all of this igneous activity be related directly to subduction of oceanic lithosphere; much of it occurred by melting of (presumably thickened) continental crust far from the ancient active margin that marked the southern edge of the European continent before its collision with Africa¹¹.

Anatectic granite also characterizes the late Cenozoic igneous intrusions in the Himalaya¹², in at least one locality in northern Tibet¹³, and in the Cordillera Oriental of the Andes^{14,15}. The geochemical signatures of each of these sequences imply melting of thickened continental crust. Moreover, high heat flow in southernmost Tibet, south of the suture zone, implies that such igneous processes continue¹⁶.

Folding and thrust faulting diminished in importance and probably ceased in most of southern Europe some time in the Carboniferous period^{3,5}. Following the cessation of thrust faulting, if not concurrent with it, roughly ENE-trending linear basins developed along at least part of the Hercynian chain^{17–20} (Fig. 1). Thick Carboniferous and Permian sedimentary rock was deposited subaerially on basement already consolidated and deeply eroded by late Palaeozoic time^{17,19,20}. Between these basins, the basement is now exposed, or younger material has been deposited directly on it. Although in some regions, Mesozoic or Cenozoic erosion could have removed the late Palaeozoic cover, elsewhere the onlap of these sequences onto crystalline rock demonstrates that in late Palaeozoic time deep basins were separated by topographically higher regions. Normal faults can be mapped on the margins of only some of the thick sequences of late Palaeozoic sedimentary rock^{18,19} (Fig. 2), but in many cases the evidence for these narrow basins being fault-bounded is simply the existence of abrupt variations in stratigraphic thickness. These basins can be recognized not only in the region shown in Fig. 1, but also over much of southern France¹⁷, and in areas highly tectonized during the formation of the Alps. The Briançonnais domain of the Franco-Italian Alps, for instance, reveals a particularly thick Triassic sequence¹⁹, associated with a well developed Permo-Carboniferous basin^{21,22}. These basins might represent only a snapshot of a phase of extension that lasted tens of million years, for extension could have begun well before the extant basins formed, with erosion destroying the record of it. In any case, normal faulting, possibly in association with strike-slip faulting, seems to have been a prevalent style of deformation in late Palaeozoic time¹⁷.

The roughly parallel grabens and half-grabens linked by strike-slip faults in the Basin and Range province of western North America (Fig. 1), in western Turkey and the Aegean, in the Tibetan Plateau, and seen less clearly in the High Andes, seem to be likely modern analogues for the landscape of the Permo-Carboniferous Hercynian belt. The evidence for active and Late Cenozoic normal faulting in Tibet^{23,24} and the Andes^{25,26} is inescapable, and that in the Basin and Range province is renowned. The Basin and Range province seems to represent a more advanced stage of crustal extension and thinning than Tibet or the Andes^{27,28}, where normal faulting may have begun only a few million years ago^{23,26}. Because as much as 60–100% crustal extension may have occurred in the Basin and Range province^{29–31}, it is very likely that in that area of Mesozoic crustal shortening the crust was thick before being thinned during Cenozoic time^{27,28,32,33}. Thus, although the analogy between the Hercynian chain and Tibet or the Andes might be imperfect if extension occurred for tens of millions of years in Europe in Palaeozoic time, the protracted extension in

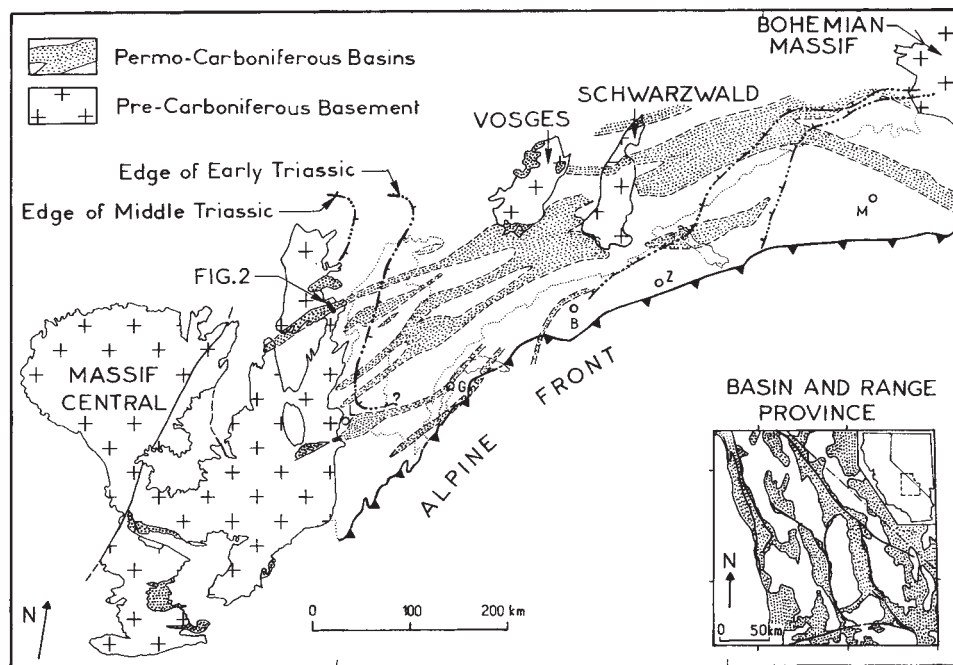


Fig. 1 Maps at the same scale of Permo-Carboniferous basins in central Europe and of the Basin and Range province in eastern California (inset), with basins dotted. Note how outcrops and subcrops^{20,34} of Permo-Carboniferous sedimentary rock define linear belts separated by pre-Carboniferous crystalline rock, and how the Triassic sedimentary rock is most extensive where the basins are most prevalent. M, München; Z, Zürich; B, Berne; G, Genève; L, Lyon.

the Basin and Range province would maintain the analogy asserted here.

During the Triassic period, portions of the Hercynian belt became submerged for the first time after crustal shortening had begun. In the area shown in Fig. 1, the classic Triassic sequence of sandstone, followed by platform limestone, and finally shale interbedded with evaporites, was deposited over the area where Permo-Carboniferous rifting seems to have been particularly active^{19,34} (Fig. 1). Correspondingly, the absence of the Triassic sequence where Permo-Carboniferous basins are sparse implies an association of the submergence with crustal extension, but the relatively smooth variations in thickness of the Triassic sequence imply that the tectonic processes that had created individual basins had ceased by Triassic time. Finally, the shallow, transient marine environments and the thickness of only 1,000–2,000 m of material deposited in 40 Myr suggest that submergence was gradual, and hence probably the result of cooling of thinned lithosphere and its underlying asthenosphere³⁵.

Beneath most of the Hercynian belt, the present crustal thickness is less than 35 km^{36–40}. In some areas, including where the Permo-Carboniferous basins are covered by Triassic sedimentary rock, the depth of the Moho is less than 30 km^{38,39}, and the basement is covered by as much as 1 or 2 km of Mesozoic marine sedimentary rock^{19,34,37}. Thus, the crystalline basement is particularly thin. We cannot eliminate Oligocene crustal thinning

as a cause of this especially thin crust, but the crustal thickness of <32 km throughout most of the Hercynian belt of France^{39,40} suggests that crustal thinning associated with late Palaeozoic extension may have left this region with crust thinner than the normal isostatically balanced 35 km.

Despite a mean elevation of ~1,500 m, the crustal thickness of the Basin and Range province is <30 km^{41,42}. At the same time, the high heat flow⁴³ and the low seismic wave velocities in the underlying mantle^{44,45} imply that the temperature within the crust and upper mantle of the Basin and Range province is higher than normal, presumably because the thickened lithospheric root detached after crustal thickening had occurred^{33,46}. Accordingly, when the upper mantle cools, the Basin and Range province should also become submerged.

Thus, by analogy with belts of active crustal extension, an initially broad, high Hercynian chain, with a deep crustal root, probably collapsed and subsided, primarily by crustal extension and crustal thinning, and later by thermally induced subsidence. The initial phase of extension would correspond to the normal faulting that characterizes the Tibetan Plateau and parts of the Andes, where the crust is still very thick. Continued extension could have thinned the crust to <35 km, as seems to have occurred in the Basin and Range province. Cooling and thermal contraction of a hot upper mantle, like that currently underlying the Basin and Range province, later would have brought the tectonically inactive late Palaeozoic Hercynian chain below sea level in early Mesozoic time. This evolution and the suggestion of a possibly similar collapse of part of the Caledonian belt in Palaeozoic time⁴⁷ are further reminders that some mountain belts, or parts of them, may be destroyed as rapidly by tectonic processes as by erosion.

This work was supported in part by NSF.

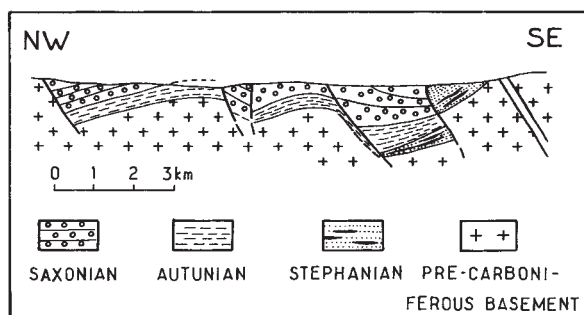


Fig. 2 Cross-section across the Permo-Carboniferous Creusot basin in France¹⁹. Note the presence of normal faults, some of which have been reactivated as reverse faults, bounding Carboniferous and pre-Carboniferous rock.

Received 15 March; accepted 6 June 1988.

- Dewey, J. F. & Burke, K. C. A. *J. Geol.* **81**, 683–692 (1973).
- Arthaud, F. *Ser. Geol. Struct. No. 1* (Pub. Univ. Sci. tech. Languedoc, France, 1970).
- Burg, J. P. & Matte, P. Z. *dt. geol. Ges.* **129**, 429–460 (1978).
- Mattauer, M. & Etchecopar, A. in *Colloque International du CNRS 268, Himalaya: Science de la Terre* 261–267 (Edns du Centre National de la Recherche Scientifique, Paris, 1977).
- Matte, P. *Tectonophysics* **126**, 329–374 (1986).
- Cazes, M. *et al. Nature* **323**, 144–147 (1986).
- Romanowicz, B. *J. geophys. Res.* **87**, 6865–6883 (1982).
- James, D. E. *J. geophys. Res.* **76**, 3246–3271 (1971).
- Chang, C. *et al. Nature* **323**, 501–507 (1986).
- Suárez, G., Molnar, P. & Burchfiel, B. C. *J. geophys. Res.* **88**, 10403–10428 (1983).
- Vidal, P. *et al. Phys. Earth planet. Int.* **35**, 179–190 (1984).

12. Le Fort, P. *J. geophys. Res.* **86**, 10545–10568 (1981).
13. Molnar, P. *et al. Science* **235**, 299–305 (1987).
14. Evernden, J. F., Kriz, S. J. & Cherroni, M. C. *Econ. Geol.* **72**, 1042–1061 (1977).
15. Tilton, G. R., Pollak, R. J., Clark, A. H. & Robertson, R. C. *R. Geol. Soc. Amer. Mem.* **154**, 791–816 (1981).
16. Jaupart, C., Francheteau, J. & Shen Xian-jie *Geophys. J. R. astr. Soc.* **81**, 131–155 (1985).
17. Arthaud, F. & Matte, P. *Geol. Soc. Am. Bull.* **88**, 1305–1320 (1977).
18. Maistre, J. de *Revue de l'Industrie minière* **45**, 541–600 (1963).
19. *Synthèse géologique du sud-est de la France, Mémoire du BRGM* (1984).
20. Rigassi, D. *Paleolab News* **2** (1977).
21. Fabre, J. *Mém. du BRGM* (1985).
22. Ellenberger, F. *Mém. du BRGM* (1954).
23. Armijo, R., Tapponnier, P., Mercier, J. L. & Han, T. *J. geophys. Res.* **91**, 13803–13872 (1986).
24. Molnar, P. & Tapponnier, P. *J. geophys. Res.* **83**, 5361–5375 (1978).
25. Dalmayrac, B. & Molnar, P. *Earth planet. sci. Lett.* **55**, 473–481 (1981).
26. Sebrier, M., Mercier, J. L., Mégard, F., Laubacher, G. & Carey-Gailhardis, E. *Tectonics* **4**, 739–780 (1985).
27. Molnar, P. & Chen, W.-P. *J. geophys. Res.* **88**, 1180–1196 (1983).
28. Molnar, P. & Lyon-Caen, H. in *Processes in Continental Lithospheric Deformation, Geol. Soc. Amer. Spec. Pap.* **218** 179–207 (1988).
29. Davis, G. A. & Burchfiel, B. C. *Geol. Soc. Am. Bull.* **84**, 1407–1423 (1973).
30. Proffett, J. M. *Geol. Soc. Am. Bull.* **88**, 247–266 (1977).
31. Wernicke, B., Spencer, J. E., Burchfiel, B. C. & Guth, P. L. *Geology* **10**, 499–502 (1982).
32. Coney, P. J. & Harms, T. A. *Geology* **12**, 550–554 (1984).
33. Sonder, L. J., England, P. C., Wernicke, B. P. & Christiansen, R. L. in *Continental Extensional Tectonics Geol. Soc. Spec. Publ.* **28** (ed. Coward, M. P., Dewey, J. F. & Hancock, P. L.) 187–201 (1987).
34. Bachmann, G. H., Müller, G. & Weggen, K. *Tectonophysics* **137**, 77–92 (1987).
35. McKenzie, D. P. *Earth planet. Sci. Lett.* **40**, 25–32 (1978).
36. Geiss, E. *Ann. Geophysicae* **5**, 623–630 (1987).
37. Meissner, R., Wever, T. & Flüß, E. R. *Ann. Geophysicae* **5**, 357–364 (1987).
38. Finckh, P., Ansorge, J., Mueller, S. & Sprecher, C. *Tectonophysics* **109**, 1–14 (1984).
39. Sapin, M. & Hirn, A. *Ann. Géophys.* **30**, 181–202 (1974).
40. Sapin, M. & Prodehl, C. *Ann. Géophys.* **29**, 127–145 (1973).
41. Prodehl, C. *Geol. Surv. Prof. Pap.* **1034** (US Govt. Print. Off., Washington, DC, 1979).
42. Smith, R. B. *Geol. Soc. Amer. Mem.* **152**, 111–144 (1978).
43. Lachenbruch, A. H. & Sass, J. H. *Geol. Soc. Amer. Mem.* **152**, 209–250 (1978).
44. Black, P. R. & Braille, L. W. *J. geophys. Res.* **87**, 10557–10568 (1982).
45. Grand, S. & Helmerberger, D. V. *Geophys. J. R. Astr. Soc.* **76**, 399–438 (1984).
46. Houseman, G. A., McKenzie, D. P. & Molnar, P. *J. geophys. Res.* **86**, 6115–6132 (1981).
47. McClay, K. R., Norton, M. G., Coney, P. & Davis, G. H. *Nature* **323**, 147–149 (1986).

Numerical simulations of thermal-chemical instabilities at the core–mantle boundary

Ulrich Hansen* & David A. Yuen†

* Institut für Geophysik und Meteorologie, Universität zu Köln, D-5000 Köln 41, FRG

† Minnesota Supercomputer Institute and Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota 55455, USA

Dynamical effects at the core–mantle boundary have so far been modelled mainly with thermal convection^{1,2}, yet accumulating evidence^{3,4} supports the idea of a combined thermal and chemical boundary layer as the likely explanation of the D'' zone. Here we present numerical simulations of thermal-chemical instabilities in the D'' layer which show that strong lateral heterogeneities in the composition and density fields can be initiated and maintained dynamically if there is continuous replenishment of material from subducted slabs coming from the upper mantle. These chemical instabilities have a tendency to migrate laterally and may help to support core–mantle boundary topography with short and long wavelengths. The thermal-chemical flows produce a relatively stagnant D'' layer with strong lateral and temporal variations in basal heat flux, which gives rise to thermal core–mantle interactions⁵, influencing the geodynamo.

The seismically resolved anomalous zone above the core–mantle boundary (CMB) holds the key to understanding the dynamical behaviour of the D'' region. Three-dimensional seismic studies^{4,6} indicate the existence of long-wavelength CMB topography. To date, most attempts^{2,7} to explain these topographic distortions have involved stresses induced by thermal convection, although more dense chemical material has also been incorporated as tracers in a thermal convection study⁸.

There are, however, some dynamical difficulties with the thermal convective mechanism. The expected decrease of the ther-

mal coefficient of expansion with depth⁹ and the increase of the lattice conductivity with pressure¹⁰ would both lead to a decrease in the thermally induced topography by several times¹¹.

Evidence^{1,4} from seismological studies indicates that the simple model of the D'' layer as a major thermal boundary layer may not be valid. The existence of a layer of dense material at the base of the mantle could provide a good explanation for the structure of the D'' layer. Such a layer might be a relic left over from core formation, or might be continuously renewed by deep subduction¹².

The dynamical consequences of the existence of both types of chemical boundary layers are explored in this work from the point of view of double-diffusive convection (DDC), which includes the participation of chemical and thermal buoyant forces. For the Earth's lower mantle the two-dimensional, cartesian DDC equations for a Boussinesq, constant-property fluid at infinite Prandtl number are described by the biharmonic momentum equation in terms of streamfunction Ψ and the evolutionary equations portraying the thermal and chemical histories:

$$\nabla^4 \Psi = Ra \frac{\partial T}{\partial x} - R_c \frac{\partial C}{\partial x} \quad (1)$$

$$\frac{\partial T}{\partial t} = \nabla^2 T + \frac{\partial \Psi}{\partial x} \frac{\partial T}{\partial z} - \frac{\partial \Psi}{\partial z} \frac{\partial T}{\partial x} \quad (2)$$

$$\frac{\partial C}{\partial t} = \frac{1}{Le} \nabla^2 C + \frac{\partial \Psi}{\partial x} \frac{\partial C}{\partial z} - \frac{\partial \Psi}{\partial z} \frac{\partial C}{\partial x} \quad (3)$$

We use a linearized equation of state for the density ρ given by $\rho = \rho_0 (1 - \alpha(T - T_0) + \beta(C - C_0))$. The horizontal coordinates are denoted by x and z with the gravitational vector aligned opposite to the z -axis. The composition of the chemical anomaly in the lower mantle is given by C , with $C = 1$ referring to 100% of chemically distinct material above the CMB. Temperature is denoted by T and the dimensionless time by t . The coefficient of thermal expansion and its analogous compositional counterpart are given respectively by α and β , with the background values denoted by the subscript 0.

We have non-dimensionalized the length by the depth d of the mantle. The Lewis number Le is given by κ/D , where κ and D are respectively the thermal and the mass diffusivities. To characterize the thermal-chemical flow we define the Rayleigh numbers by $Ra = \alpha g \Delta T d^3 / \kappa \nu$ and $R_c = \beta g \Delta C d^3 / \kappa \nu$, where ν , ΔT and ΔC refer to the kinematic viscosity, the temperature drop across the mantle and the compositional difference across the D'' layer, respectively. The ratio of chemical to thermal buoyancy is given by $R_p = R_c / Ra$. R_p is an important parameter in this problem: a value greater than zero represents a situation of dense material in the D'' layer. Values of R_p are still relatively unconstrained by seismology and by our meagre knowledge of the equation of state. We have adopted this continuum or field approach of modelling thermal-chemical convective flows instead of using discrete tracers^{8,13} to represent distinct chemical species because of its greater simplicity and also accuracy in the local statistical interpretation of the compositional and density fields.

The calculations were made with a finite-element method¹⁴. For details of the grid configuration refer to Fig. 1. The CMB is represented by a stress-free, and isothermal boundary with either the composition C set to unity or $\partial C / \partial z = 0$ (no mass flux). Geophysically, the first situation implies a supply of hot, enriched material by treating Le as an effective diffusion coefficient that parameterizes a complex physical process. The non-flux boundary condition simulates the existence of a residual layer that has been imposed in the beginning.

We will focus our attention on the interaction between the large-scale circulation and the thermal-chemical instabilities developed at the basal boundary layer. Most of the calculations were done in a rectangular domain with an aspect ratio of 1.8

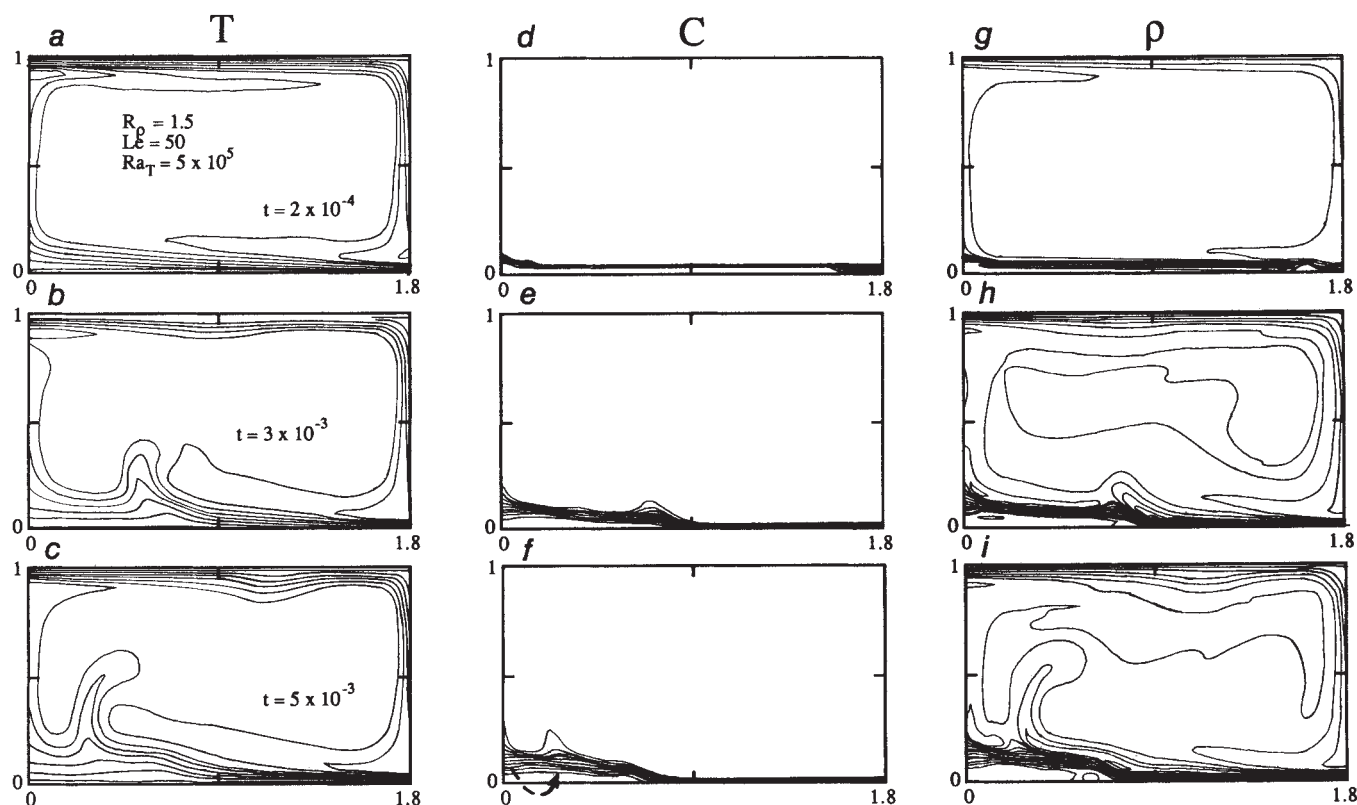


Fig. 1 Sequence of snapshots showing the development of the temperature T , composition C and perturbed density ρ -fields. Perturbed density is defined as the deviation from the background density due to thermal and chemical fluctuations. The Rayleigh number Ra is 5×10^5 , $R_p = 1.5$ and $Le = 50$. Calculations are done over a 45×45 non-uniform mesh with grid refinement by a factor of ten just above the CMB. Timestepping is performed by a second-order predictor corrector method with a timestep less than given by the Courant-Friedrichs-Levy criterion. Arrows in f indicate schematically the secondary circulation within the thermal and chemical boundary layer. Time has been nondimensionalized by d^2/κ . The bottom boundary is kept on a fixed concentration $C = 1$. Thermal overturn time is about 0.01.

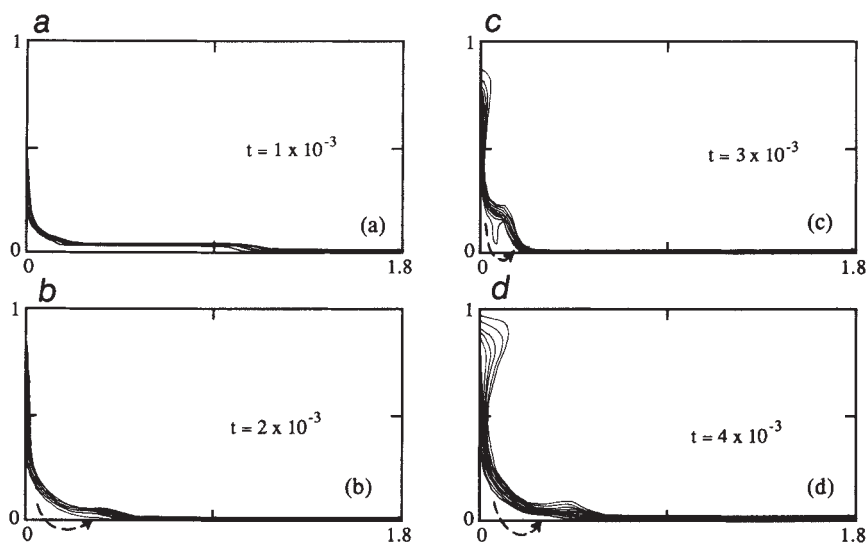
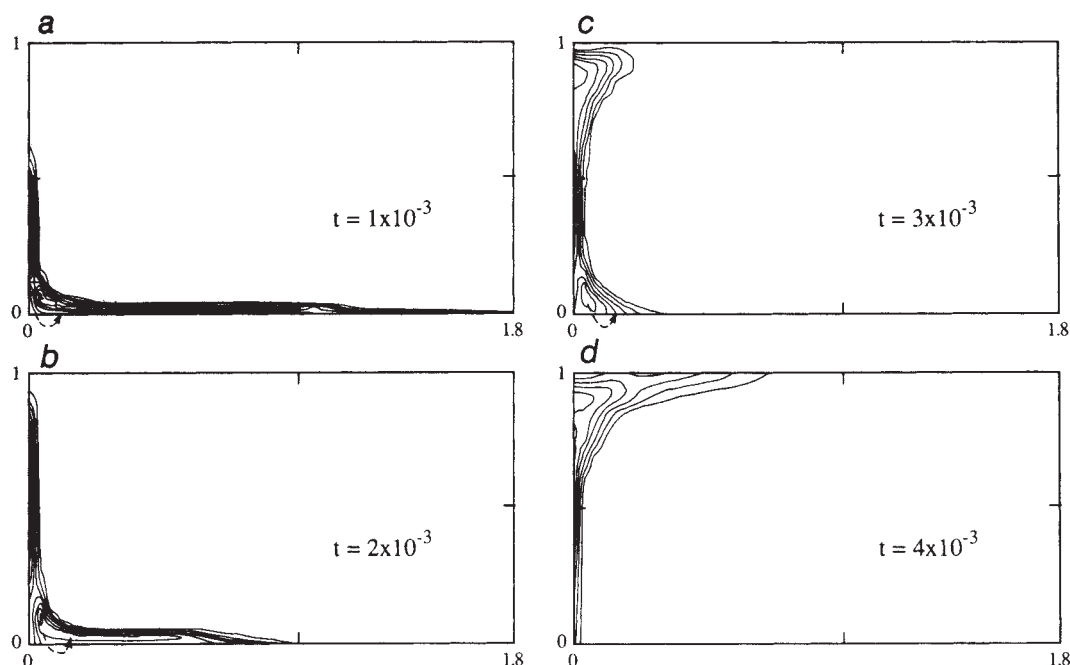


Fig. 2 Evolution of the compositional field C for $R_p = 0.5$. Ra , Le and boundary conditions are the same as in Fig. 1. Arrows indicate secondary circulations. Thermal overturn time is about 0.01.

because of the ease of inducing local boundary layer instabilities in a global circulation¹⁵. The initial temperature field was taken from a solution of the steady-state thermal convection at the appropriate Ra . We impose at the bottom a chemical layer of thickness a , which is an important controlling variable in this problem. In what follows in Figs 1 and 2, the initial thickness a has to be set to 0.03 (90 km). In this regard, Davies and Gurnis⁸ also found it necessary to impose a chemical thickness of 0 (100 km).

Figure 1 shows a sequence of temperature, compositional and density fields illustrating the development of thermal-chemical instabilities in D'' with $Ra = 5 \times 10^5$, $R_p = 1.5$, $Le = 50$ and a fixed concentration at the bottom. First we observe that the bottom boundary layer is immediately thickened relative to that obtained in pure thermal convection. The next stage (Fig. 1b, e and h) is characterized by the formation of compositional heterogeneities (Fig. 1e). A compositional protuberance, which was formed earlier at the right, starts moving to the left. The

Fig. 3 Evolution of the compositional field C for same conditions as in Fig. 2, but with a no-flux condition for C at the bottom ($\partial C/\partial z = 0$). Thermal overturn time is about 0.01.



compositional bumps at the CMB move because of the presence of the large-scale circulation. These hill-like features are clearly of a dynamical origin. There occurs at the same time a complex interaction between the cold descending and a growing boundary layer instability emerging from the bottom (Fig. 1b). Increasing amounts of lateral variation in the bottom heat flux are developed as a consequence of the presence of compositional heterogeneities. The bottom row (Fig. 1c, f and i) shows a later stage, where the lateral differences in the compositional and density fields become more accentuated and compressed. The averaged bottom heat flux decreases with time, as the D'' layer becomes convectively more stagnant because of the accumulation of chemical heterogeneities.

Secondary currents are produced inside the chemical boundary layers (see arrows in Fig. 1f) in which heavy material is being drawn up by the thermal instability. We note that both short- and long-wavelength density heterogeneities can be generated at the CMB from thermal-chemical instabilities. The amplitude of the compositional heterogeneities grows in step with time and should closely reflect, at each instant, the CMB topography⁸. For this value of R_p ($R_p = 1.5$), not much upward leaking of the enriched material (Fig. 1e and f) takes place.

Figure 2 shows the evolution of the compositional field for a smaller amount of chemical buoyancy ($R_p = 0.5$). It is seen that shorter wavelength chemical heterogeneities are developed of larger amplitude than those with $R_p = 1.5$. There is also an upward leaking of enriched material through the ascending plume, driven by thermal buoyancy. Secondary circulations develop inside the chemical boundary layer (see arrows in Fig. 2), as lighter material is entrained downward by the counter-rotating flow (Fig. 2c). The wavelength of the heterogeneities is subsequently increased by this secondary flow (Fig. 2d) and also by the push imparted by the laterally moving 'tongue' (see Fig. 1c as an example).

We have found that increasing Le from 50 to 300 does not alter the behaviour of the flow, as long as the initial thickness is kept constant at 90 km. But additional numerical experiments show that for an extremely thin chemical layer, which is probable at higher values of Le , lateral heterogeneities cannot be generated, at least in a short time-scale. We conclude this on the basis of calculations made with the same parameters as in Fig. 1 but using an initial thickness set to 0.01 d (30 km). Thus solid-state diffusion processes, with characteristically large values of Le of

at least 10^8 , cannot play any significant role in maintaining chemical heterogeneities at the CMB.

To investigate the importance of a continuous renewal we calculated the development of a scenario using the same parameters as given in Fig. 2, but with no-flux conditions for the composition at the bottom. This situation may also mimic the case with an extremely high Le , where diffusion is minimal at the CMB. Figure 3 shows that initially a heterogeneity develops (Fig. 3a, b). Subsequently the chemical layer loses its density by leaking and internal convection and is advected up and away (Fig. 3c, d). We then conduct a systematic study of both types of boundary conditions for $Ra = 10^4$ and an aspect ratio of one. This investigation confirmed the result shown in Fig. 3 that, for no-flux conditions, a higher R_p is necessary to maintain a persistent layer of dense material. At Le varying between 30 and 300, $R_p > 2$ is required for the no-flux case, whereas $R_p \sim 1$ is sufficient with an isochemical boundary. For $Le = 100$ and $Ra = 10^4$, we have found that the time t_R , defined here to be the time in which the maximum concentration decreases to 1/10 of the initial maximal value in no-flux situations, increases linearly with R_p for R_p between 3 and 10. These values of R_p are reasonable, if one takes the decreasing values of thermal expansivity with depth into account⁹. This relationship between t_R and R_p shows the stabilizing influence of chemical buoyancy. Previous investigation⁸ using chemical tracers did not quantify this basic difference in the dynamical behaviour arising from the boundary conditions for the compositional field at the CMB.

If the D'' layer at the base of the mantle is a combined thermal and chemical boundary layer for suitable parameter values, then instability in layers with sufficiently high values of R_p and suitable thickness of the chemical layer would lead to a convectively sluggish D'' layer, and would produce strong chemical and density heterogeneities both laterally and vertically. In the case of the chemical layer not being renewed, either a sufficiently high density contrast or a reduced thermal expansivity ($R_p > 2$), or a sufficiently thick initial layer ($a > 90$ km) would be needed for the layer to survive over geological epochs. The value for the thickness of the layer from seismological evidences³ might be slightly high and might not be favourable under the expected mantle conditions. Regarding the possible renewal process, deep subduction of the crust may be a viable mechanism for accumulating dense substances ($R_p > 0$) at the CMB. On the other hand, chemical reactions at the CMB can potentially be a source of

material with high R_p (iron enrichment), but only if some amounts of partial melting of the CMB from a hot core¹⁶ are present. In this case, the effective mass diffusivity may be sufficiently enhanced by transport mechanisms involving two-phase flow in melts¹⁷, which is much more efficient than solid-state diffusion.

The development of the laterally moving thermal-chemical heterogeneities may lead to topographic variations with both short- and long-wavelength features. Dynamically induced amplitudes of CMB topography of several kilometres have been found in the chemical-tracer model⁸. The thermal-chemical instabilities, discussed above, may offer a way to explain the unsupported CMB topography, which is difficult to explain simply on the basis of pure thermal convection⁷. In addition to providing topography, compositional variations can produce modulations of the bottom heat flux, more so than is found in ordinary thermal convection. Such spatial and temporal modulations of the bottom heat flux would certainly have important ramifications on the thermal coupling between mantle and core processes and on time-dependent dynamics of core convection.

We thank Raymond Jeanloz and IGPP Mantle Convection Workshop (Los Alamos, NM), Uli Christensen, Volker Steinbach and David Stevenson for stimulating discussions on this subject. This work was supported by the American National Science Foundation, by NASA and by the Deutsche Forschungsgemeinschaft. We appreciate the help of Kari L. Rabie in preparing the manuscript.

Received 5 January; accepted 1 June 1988.

- Hager, B. H., Clayton, R. W., Richards, M. A., Comer, R. P. & Dziewonski, A. M. *Nature* **313**, 541–545 (1985).
- Zhang, S. & Yuen, D. A. *Geophys. Res. Lett.* **14**, 899–902 (1987).
- Young, C. J. & Lay, T. A. *Rev. Earth planet. Sci.* **15**, 25–46 (1987).
- Creager, K. C. & Jordan, T. H. *Geophys. Res. Lett.* **13**, 1497–1500 (1986).
- Bloxham, J. & Gubbins, D. *Nature* **325**, 511–513 (1987).
- Morelli, A. & Dziewonski, A. M. *Nature* **325**, 678–683 (1987).
- Olson, P., Schubert, G. & Anderson, C. *Nature* **327**, 409–413 (1987).
- Davies, G. F. & Gurnis, M. *Geophys. Res. Lett.* **13**, 1517–1520 (1986).
- Anderson, D. L. *Phys. Earth planet. Inter.* **45**, 307–323 (1987).
- Brown, J. M. *Geophys. Res. Lett.* **13**, 1509–1512 (1986).
- Zhang, S. & Yuen, D. A. *Geophys. Res. Lett.* **15**, 451–454 (1988).
- Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421–436 (1982).
- Christensen, U. R. *EOS* **68**, 1488 (1987).
- Hansen, U. & Yuen, D. A. *Geophys. Res. Lett.* **14**, 1099–1102 (1987).
- Hansen, U. & Ebel, A. *Geophys. J.* **194**, 181–191 (1988).
- Williams, Q., Jeanloz, R., Bass, J., Svendsen, B. & Ahrens, T. J. *Science* **236**, 181–182 (1987).
- McKenzie, D. P. *J. Petrology* **25**, 713–765 (1984).

Effects of long-term variation on the frequency spectrum of the geomagnetic reversal record

Timothy M. Lutz* & Geoffrey S. Watson†

* Department of Geology, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6316, USA

† Department of Statistics, Princeton University, Princeton, New Jersey 08544, USA

Several studies have suggested a period of about 30 Myr in the frequency of geomagnetic reversal^{1–4} superposed on a long-term variation, but none of these studies explicitly investigated the effects of the long-term variation on their analyses. We show that the long-term variations alone account for most of the spectral evidence taken to support the 30 Myr periodicity. In particular, high peaks in the spectra near 30 Myr, the overall pattern of peaks, and shifts of the peak pattern for subsets of the data are explained well by a long-term variation model. We suggest that there may be periodic changes in reversal frequency about the long-term variation, but that more study will be required before the periodicity can be established.

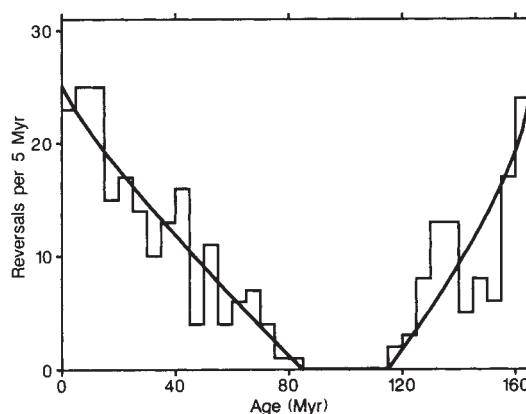


Fig. 1 The step function shows the number of magnetic reversals per 5 Myr in the Harland *et al.*⁵ record. The smooth curve is the aperiodic intensity function, defined in equation (1), used to model the long-term variations in frequency.

Our procedure is to model the long-term variations in the Harland *et al.*⁵ magnetic reversal record by an intensity function (the time rate, or frequency, of reversals). Our model function approximates the high reversal frequencies near the present and decreases to zero during the Cretaceous quiet superchron (Fig. 1). The integral of the function over the last 85 Myr yields the observed number of reversals (196). The function approximates the increase in reversal frequency from 115 to 165 Myr ago, and its integral over that range yields the observed number of reversals (99). The model function represents the long-term variations but does not exhibit any periodicity. Its formula is

$$\lambda(t) = 25.09 - 0.575t^{0.85} \quad 0 \leq t \leq 85$$

$$\lambda(t) = 0 \quad 85 \leq t \leq 115$$

$$\lambda(t) = 24.04 - 4.797 \left(\frac{165-t}{5} \right)^{0.70} \quad 115 \leq t \leq 165 \quad (1)$$

We show the response of Fourier analysis to the model intensity function by using numerical integration to construct a spectrum of $\lambda(t)$ with the formula

$$R_p(\tau_1, \tau_2) = \left[\left(\int_{\tau_1}^{\tau_2} \lambda(t) \sin \left(\frac{2\pi t}{p} \right) dt \right)^2 + \left(\int_{\tau_1}^{\tau_2} \lambda(t) \cos \left(\frac{2\pi t}{p} \right) dt \right)^2 \right]^{1/2} \quad (2)$$

Here $\lambda(t)$ is the model intensity function, t is the time and p is the period. Equation (2) allows us to use the model over any interval (τ_1, τ_2) in $(0, 165)$.

The spectrum of the actual reversal record (called circular dispersion analysis by Lutz⁶) is given by

$$R_p = \left[\left(N^{-1} \sum \sin \frac{2\pi t_i}{p} \right)^2 + \left(N^{-1} \sum \cos \frac{2\pi t_i}{p} \right)^2 \right]^{1/2} \quad (3)$$

where the t_i are the times of the N actual reversals and p is a trial period. Comparison of the spectrum of $\lambda(t)$ (Fig. 2a) and the spectrum of the data (Fig. 2b) reveals the broad similarities between them. The positions and widths of the peaks are virtually identical.

We also use the model intensity function to simulate the magnetic reversal data. The simulated reversal records are generated as follows: if we define $\lambda(t) dt = dT(t)$ for $0 \leq t \leq t_0 = 165$ Myr, then $T(t) = \int_0^t \lambda(t') dt'$ increases with t from 0 to $T(t_0)$ = total number of reversals in the record. T is a new timescale, and

$$\lambda(t) dt = \text{Prob (a reversal in } (t, t+dt))$$

$$= dT = \text{Prob (a reversal in } (T, T+dT))$$

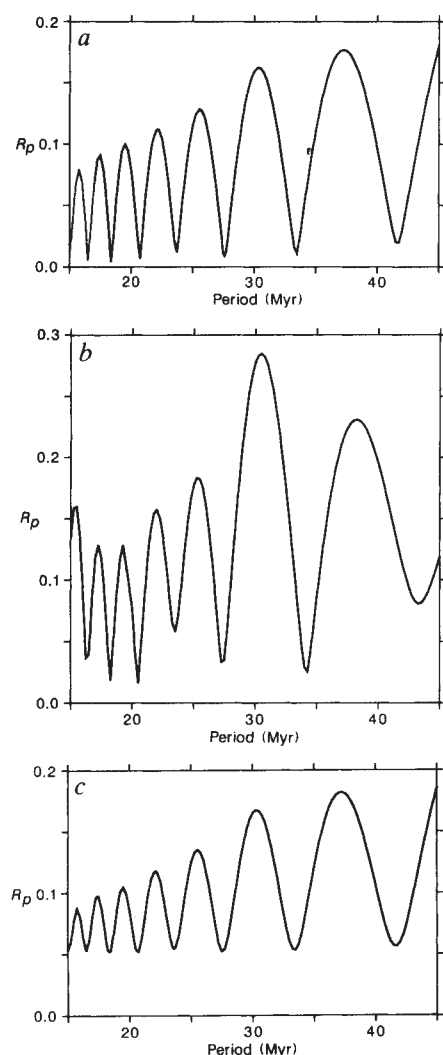


Fig. 2 *a*, The Fourier spectrum of the model intensity function over the range of 0 to 165 Myr. R_p is defined in equation (2). *b*, The spectrum of the magnetic reversal record from 0 to 165 Myr. R_p is defined in equation (3). *c*, An average spectrum constructed by superimposing the spectra of 500 simulations of the magnetic reversal record.

Thus, the occurrence of events in the T timescale is very simple. It is the well known Poisson process (see ref. 7 for example) and trivial to simulate. The simulated T -times are transformed back into t -times by solving $T(t) = T$ for t .

The average of spectra of 500 simulations based on $\lambda(t)$ shows the same pattern of peaks (Fig. 2c) as in Fig. 2a and b. By averaging these spectra, the peaks are reduced and the troughs filled in, leading to a smaller amplitude of variation. Apparently, realizations based on the aperiodic long-term model yield a pattern of peaks similar to that obtained for the actual data.

Furthermore, the location of the highest spectral peak for these simulations is non-uniform. It occurs at $30 \text{ Myr} \pm 1 \text{ Myr}$ in over 28% of the simulations (Fig. 3) for periods between 15 Myr and 45 Myr. Thus, there is a substantial chance that a simulated set of data will yield a spectrum with its largest peak at 30 Myr and have the same peak distribution pattern as the actual reversal record.

Stothers⁴ called attention to the way the pattern of peaks changes if subsets of the data are analysed, concluding that the changes in the peak pattern, together with the persistence of the highest peak near 30 Myr for all subsets, are strong evidence for a periodicity. We have computed the spectrum (2) on ranges (τ_1, τ_2) of the model intensity function that duplicate the subsets of data analysed by Stothers⁴ and find that the peak patterns

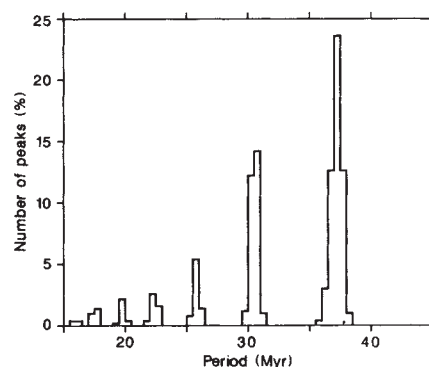


Fig. 3 The distribution of the period at which the highest peak occurred for the 500 simulations used to compute the average spectrum in Fig. 2c. The ordinate shows the number of peaks (expressed as a percentage of 500) with a period within cells with a width of 0.5 Myr. The sum of the ordinates of the cells within 1 Myr of 30 Myr is 28%.

are virtually identical (Table 1).

An examination of Fig. 1 (and also Fig. 1 in ref. 2, in which the bins have been shifted) suggests that there could be a 30 Myr variation in reversal frequency superimposed on the smooth long-term variations. Raup², using McFadden and Merrill's suggestion⁸ of simplified long-term variation, claimed to have found evidence that the observed 30 Myr peak is significantly high, adopting the view that a 30 Myr periodicity was the most simple interpretation of his results. However, peak-height tests do not unequivocally indicate periodicities because significantly high peaks can result from a combination of long-term and shorter-term variations. We have shown that more than one explanation may be required, more subtle methods being needed to uncouple the effects of long-term variation from shorter-term periodicities (to be considered elsewhere).

We also query whether 'periodicity' is a constructive term for referring to data that depart significantly from some null model. Lutz⁹ has shown that the ages of geological events, such as mass extinctions, can be explained satisfactorily by a variety of models, some periodic and some episodic. Spectral peaks exceeding 'random' expectations were shown not to require a highly periodic process, and Fourier methods were concluded as not being so good at differentiating stationary periods from other non-Poisson structures in short data sets. In this regard, it must be remembered that the magnetic reversal record would contain only 5.5 cycles of a 30 Myr periodicity. Furthermore, a quiet superchron, roughly one cycle in length and unexplained by the periodic model, separates the data into two parts, the longest of which would contain fewer than three cycles.

In conclusion, we suggest that the variations in magnetic reversal frequency over the last 165 Myr could be explained

Table 1 Periods of spectral peaks resulting from analysis of progressively truncated magnetic reversal records⁵

$\tau_1 - \tau_2$		Periods (Myr)							
0-165	A	48	38	30	25	22	19	17	16
	B	ND	37	30	26	22	19	17	16
	C	48	37	30	26	22	20	17	16
15-165	A	46	33	28	23	20		18	16
	C	44	34	28	23	20		18	16
30-165	A	41	30		24	21		18	16
	C	40	30		25	21		18	16
45-165	A	34	28		22			18	16
	C	35	27		22			19	16

A, Peaks from ref. 4, Table 1. B, Peaks in the average spectrum of simulated reversal records. C, Peaks from Fourier analysis of the model intensity function. ND, Not determined.

reasonably well by a long-term variation model. This study, along with other recent criticisms of the testing of magnetic reversal periodicity^{10,11}, suggests that previous claims that "a statistically significant period of ~30 Myr does formally exist" were premature⁴. Further investigation is necessary to show whether the height of the 30 Myr peak in the spectrum requires a short-term explanation.

Computer equipment was obtained through a grant from the IBM Advanced Education Project. G.S.W. acknowledges support from the National Science Foundation and thanks P. Bloomfield for discussion.

Received 28 December 1987; accepted 23 May 1988.

1. Negi, J. G. & Tiwari, R. K. *Geophys. Res. Lett.* **10**, 713–716 (1983).
2. Raup, D. M. *Nature* **314**, 341–343 (1985).
3. Pal, P. C. & Creer, K. M. *Nature* **320**, 148–150 (1986).
4. Stothers, R. B. *Nature* **322**, 444–446 (1986).
5. Harland, W. B. *et al.* *A Geological Time Scale* (Cambridge University Press, 1982).
6. Lutz, T. M. *Nature* **317**, 404–407 (1985).
7. Feller, W. *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1950).
8. McFadden, P. L. & Merrill, R. T. *J. geophys. Res.* **89**, 3354–3363 (1984).
9. Lutz, T. M. *Geology* **15**, 1115–1117 (1987).
10. Stigler, S. M. *Nature* **330**, 26 (1987).
11. McFadden, P. L. *Nature* **330**, 27 (1987).

A comparative approach to predicting competitive ability from plant traits

Connie L. Gaudet & Paul A. Keddy

Biology Department, University of Ottawa, 30 Somerset East, Ottawa, Ontario K1N 6N5, Canada

Decades of study of interspecific competition in community ecology has yielded an overwhelming body of special cases but few general principles^{1–3}. This is largely because of the phenomenological, non-predictive approach used⁴. Further progress requires a predictive approach⁵ that will enable general principles to be deduced that apply beyond the species and conditions of a particular study or site. General principles are best sought using a comparative approach, that is, the systematic screening of a large number of species under standardized experimental conditions^{6,7}. We used 44 wetland plant species to test whether competitive ability could be predicted from plant traits. Multiple linear regression showed that there was a strong relationship between plant traits and competitive ability ($r^2 = 0.74$). Plant biomass explained 63% of the variation in competitive ability and plant height, canopy diameter, canopy area and leaf shape explained most of the residual variation. This study represents a major step in escaping the current phenomenological approach to competition in community ecology, and provides a general predictive tool for studying competition in natural communities.

To date there has been no systematic analysis of the relationship between plant traits and experimental measures of competitive ability in a large, morphologically diverse, multi-species community. Several authors have advocated the use of physiological, morphological or behavioural traits to predict competitive ability^{1–4,8}. The present study measures the relative competitive ability of 44 herbaceous plant species and tests whether it is correlated with simple measurable plant traits.

A modified additive design^{9,10,11} was used to assess competitive ability as the relative ability of each species to suppress the growth of a common indicator species, or phytometer. All 44 species were tested against the phytometer, *Lythrum salicaria*. A subset of 10 species was tested against *Penthorum sedoides* to assess the effect of different phytometers on the results. Single phytometers were planted in the centre of a 1-litre pot in a

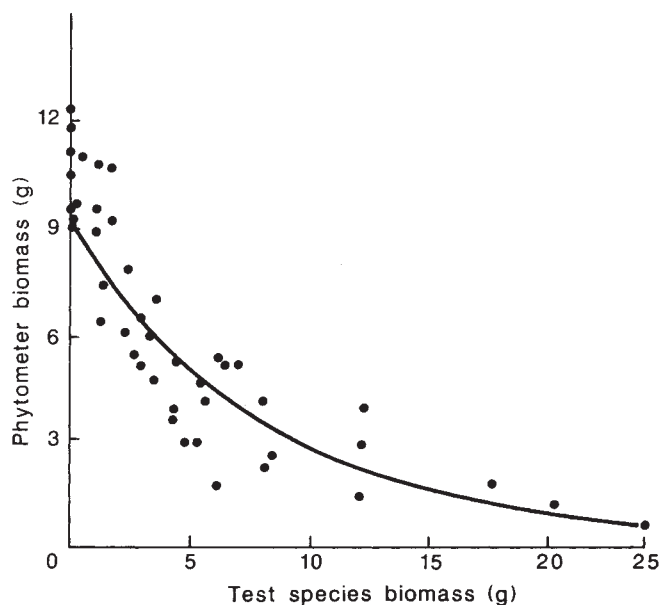


Fig. 1 Phytomer biomass (PB) (*L. salicaria*) as a function of test-species biomass (TSB) ($n = 44$). $PB = \exp(2.217 - 0.118 \text{ TSB})$; $F = 186.3$, $P < 0.0001$, $r^2 = 0.81$; $b = 0.118 \pm 0.0021$, $p < 0.05$. Test species were (in ascending order of competitive ability): 1, *Ranunculus reptans*; 2, *Lobelia dortmanna*; 3, *Xyris difformis*; 4, *Juncus militaris*; 5, *Eleocharis palustris*; 6, *Juncus filiformis*; 7, *Drosera intermedia*; 8, *Juncus pelocarpus*; 9, *Sabatia kennedyana*; 10, *Anemone canadensis*; 11, *Rhynchospora fusca*; 12, *Eriocaulon septangulare*; 13, *Panicum longifolium*; 14, *Eleocharis erythropoda*; 15, *Dulichium arundinaceum*; 16, *Onoclea sensibilis*; 17, *Lysimachia terrestris*; 18, *Leersia oryzoides*; 19, *Impatiens capensis*; 20, *Viola lanceolata*; 21, *Triadenum fraseri*; 22, *Galium palustre*; 23, *Carex crinita*; 24, *Lysimachia nummularia*; 25, *Spartina pectinata*; 26, *Scirpus validus*; 27, *Polygonum hydropiperoides*; 28, *Hypericum ellipticum*; 29, *Iris versicolor*; 30, *Mentha arvensis*; 31, *Acorus calamus*; 32, *Eupatorium maculatum*; 33, *Rumex verticillatus*; 34, *Potentilla anserina*; 35, *Lysimachia thysiflora*; 36, *Lysimachia ciliata*; 37, *Carex rostrata*; 38, *Scirpus fluviatilis*; 39, *Pilea pumila*; 40, *Typha latifolia*; 41, *Stachys palustris*; 42, *Phalaris arundinacea*; 43, *Bidens cernua*; 44, *Lythrum salicaria*. Nomenclature follows Fernald²⁸.

sterile, organic, high-nutrient mix. Four individuals of each test species were planted in a systematic design around the phytometer (five replicates for each species). Most studies on competitive interference within plant communities are based on considerations of population density, but this obscures variation due to other factors, such as plant size^{12,13}. Therefore, this study was conducted at a single density with the effects of distance and angular dispersion kept constant, so that the effects of other parameters could be clearly interpreted. Plants were grown over one growing season, April 1986 to September 1986. Several morphological measurements (Table 1) were made on the test species both under the experimental conditions described above, and on each species grown singly.

When all 12 measured variables were included, multiple regression showed a strong predictive relationship between plant traits and competitive ability ($r^2 = 0.74$). The biomass of the phytometer was correlated with several traits, but most strongly with the above-ground biomass of the test species (Table 1). Height was also highly correlated with phytometer biomass, but in a stepwise multiple regression it increased r^2 by only 2% over above-ground biomass alone. These relationships can be better described by an exponential equation (Figs 1 and 2).

A major objective was to identify general relationships between plant traits and competitive ability. We therefore conducted several other experiments to determine whether the

relationships were robust. The morphological traits discussed above were measured on plants interacting with the phytometer. To examine the influence of the phytometer on morphology, the same traits were measured on each test species grown singly, but under otherwise identical experimental conditions. When regressed against phytometer biomass, similar results were obtained with both sources of measurement and test species biomass was still most strongly correlated with competitive ability ($r^2 = 0.74$; $P < 0.001$).

Changing the phytometer has no measurable effect on the results. Results using *L. salicaria* were compared with results obtained using a smaller species, *P. sedoides*, as a phytometer. Biomass was still the most significant predictor of competitive ability ($r^2 = 0.75$; $p < 0.001$) and the slopes of the regression equations (phytometer biomass as a function of test biomass) were not significantly different (*L. salicaria*, $b = -0.118$; *P. sedoides*, $b = -0.207$; $t = 0.35$).

These results suggest that 'guilds'¹⁴ or 'functional groups'¹⁵ of plants with similar competitive ability may occur along an intergrading continuum of increasing biomass. Within these guilds, that is, where biomass is similar, other factors may become more important in determining the outcome of competitive interactions. Therefore, though our results show that biomass has the best predictive value across this range of species morphologies, in a study that considers a small number of species with a small biomass range, height or other life-history or morphological variables could be the critical determinants of dominance. An inherent assumption of replacement-series experiments⁹ is that the species are of similar biomass, so that this traditional approach to measuring competitive ability becomes an assessment of effects 'per biomass'. This may explain why such studies have failed to yield general relationships and patterns.

Though our experiment clearly establishes a predictive relationship between competitive ability and plant traits, it was conducted at a single nutrient level deemed to be most representative of the highly productive conditions of wetlands¹⁶ and those in which competition is presumably most important^{8,17}. Our results do not provide a mechanistic interpretation of resource competition. Biomass may simply integrate or summarize other traits, such as high rates of resource capture above and below ground^{18,19}. Alternatively, it has been suggested that at high nutrient levels, light becomes the critical limiting resource^{20,21} and traits we have identified may be primarily associated with competition for light.

What are the general implications of these results in understanding the organization of natural communities? Published research suggests that size-related variables are pervasive

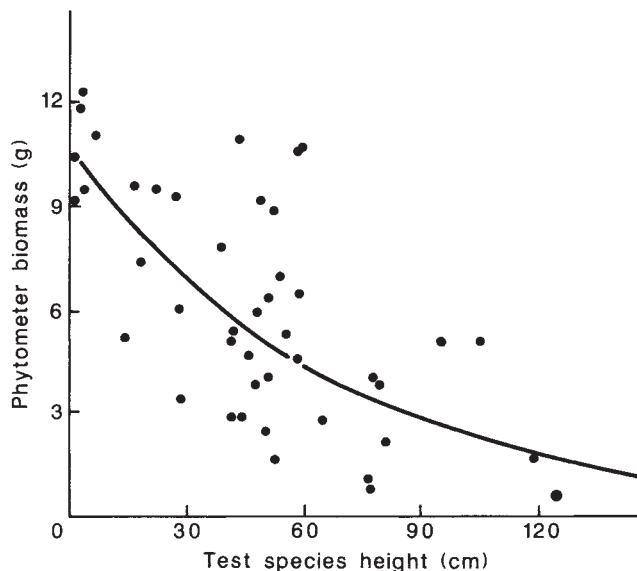


Fig. 2 Phytometer biomass (*L. salicaria*) as a function of test species height (TSH) ($n = 44$). $PB = \exp(2.392 - 0.016 \text{ TSH})$; $F = 36.263$, $P < 0.05$, $r^2 = 0.64$.

indicators of competitive dominance in a variety of herbaceous plant communities^{1,3,22-25} and within forest monocultures¹². It has been shown that plant biomass or height varies predictably along natural gradients of stress and disturbance^{26,27}. Successional and even evolutionary patterns are represented by increases in plant biomass²⁰. The occurrence of natural gradients of plant biomass in the field, and the relationship established in this study between plant traits and competitive ability, suggests that competitive ability is an important determinant of community pattern. Our results provide a general predictive tool for testing this relationship. The ultimate value of this research is that it shows the potential for predicting the outcome of multispecies competitive interactions, an important goal if community ecology is to move from a phenomenological to a predictive science.

This study was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

Received 4 March; accepted 24 May 1988.

Table 1 Correlation (r) between traits of 44 test species and biomass of the phytometer (*Lythrum salicaria*)

Plant traits	Correlation with phytometer biomass
Biomass, total (g)	-0.775**
Biomass, above ground (g)	-0.791**
Biomass, below-ground (g)	-0.710**
Height (cm)	-0.659**
Leaf length (cm)	-0.084
Leaf width (cm)	-0.179
Leaf area (cm ²)	-0.302
Leaf number	-0.244
Leaf shape (length:width)	0.356*
Canopy diameter (cm)	-0.455*
Canopy area (cm ²)	-0.593**
Shoot to root ratio (g/g)	-0.016

* $P < 0.05$; ** $P < 0.001$. Correlations are simple linear correlations with phytometer biomass.

- Keddy, P. A. *Competition* (Chapman and Hall, London, in the press).
- Grime, J. P. *Nature* **250**, 26-31 (1974).
- Grime, J. P. *Nature* **244**, 344-347 (1973).
- Tilman, D. *Am. Nat.* **129**, 769-774 (1987).
- Peters, R. H. in *Conceptual Issues in Ecology* (ed. Saarinen, E.) 215-227 (Reidel, Dordrecht, 1980).
- Grime, J. P. & Hunt, R. *J. Ecol.* **63**, 393-422 (1975).
- Grime, J. P. *Plant Strategies and Vegetation Processes* (Wiley, Chichester, 1979).
- Grime, J. P. *Am. Nat.* **111**, 1167-1194 (1977).
- Harper, J. L. *Population Biology of Plants* (Academic, New York, 1977).
- Goldberg, D. E. *Ecology* **68**, 1211-1233 (1987).
- Welbank, P. J. *Ann. appl. Biol.* **51**, 107-125 (1963).
- Weiner, J. *J. Ecol.* **72**, 185-196 (1984).
- Penridge, L. K. & Walker, J. *J. Ecol.* **74**, 925-936 (1986).
- Root, R. B. *Ecol. Monogr.* **37**, 317-350 (1967).
- Day, R. T., Keddy, P. A., McNeill, J. & Carleton, T. *Ecology* (in the press).
- Auclair, A. N. D., Bouchard, A. & Pajaczkowski, J. *J. Ecol.* **57**, 941-952 (1976).
- Wilson, S. D. & Keddy, P. A. *Ecology* **67**, 1236-1242 (1986).
- Donald, C. M. *Aust. J. agric. Res.* **9**, 421-432 (1958).
- Chapin, F. S. A. *Rev. Ecol. Syst.* **11**, 233-260 (1980).
- Tilman, D. in *Community Ecology* (eds Diamond, J. & Case, T. J.) 359-380 (Harper and Row, New York, 1986).
- Schmitt, J., Ehrhardt, D. W. & Cheo, M. *Ecology* **67**, 1502-1507 (1986).
- Mack, R. N. & Harper, J. L. *J. Ecol.* **65**, 345-363 (1977).
- Goldsmith, F. B. *J. Ecol.* **66**, 921-931 (1978).
- Snaydon, R. W. & Howe, C. D. *J. appl. Ecol.* **23**, 667-674 (1986).
- Mitchley, J. & Grubb, P. J. *J. Ecol.* **74**, 1139-1166 (1986).
- Wilson, S. D. & Keddy, P. A. *Am. Nat.* **127**, 862-869 (1986).
- Menges, E. S. & Waller, D. M. *Am. Nat.* **122**, 454-473 (1983).
- Fernald, M. L. *Gray's Manual of Botany* (Van Nostrand, New York, 1970).

Female visual displays affect the development of male song in the cowbird

Meredith J. West* & Andrew P. King†

* Department of Psychology, University of North Carolina, Chapel Hill, North Carolina 27599, USA

† Department of Psychology, Duke University, Durham, North Carolina 27706, USA

The role of social stimulation in avian vocal learning is well documented¹. The separate contribution of social, as opposed to vocal, stimulation has been difficult to address, however, because in almost all cases both tutor and pupil sing. The opportunity to isolate such effects arose in cowbirds (*Molothrus ater ater*) after discovering that males housed with non-singing female cowbirds made vocal changes which related directly to the female preferences for native song²⁻⁴. Here we report how females communicate with males about songs. We describe a visual display by females, a wing stroke, that is elicited by specific vocalizations. The songs that trigger wing strokes are in turn highly effective releasers of copulatory postures, and thus this previously unnoticed female display has biological significance. The data not only provide the first evidence of the tutorial role of male-female interactions during song ontogeny, they also clearly implicate visual stimulation in song learning, a process that has until now been assumed to be affected only by auditory information⁵.

The study was conducted in two parts. First, interactions between captive male-female pairs were videotaped in March and April, the time when cowbirds are found on their breeding habitat in the wild. Second, playbacks of songs that had elicited specific reactions from females during the videotaped sessions were presented one year later to females in breeding condition, thus testing the songs' effectiveness as releasers of copulatory postures.

We examined social interactions in eight male-female pairs in the first part of the study. The males were wild-caught juveniles captured when under 50 days of age; the females were adults captured the year before. Individual pairs were housed in sound-attenuating chambers and maintained according to previously described procedures⁶. The males were exposed to tape recordings of male songs for three months in the fall (autumn). Synchronized audio and video recordings were made in March and early April, at which point song types could be identified. These were stereotyped acoustic patterns differing from one another in the frequency contours of individual notes, note syntax, and the modulation of terminal whistles. Ten hours of singing were videotaped for four pairs and 12 hours for four other pairs. Audio taping continued into May to record the males' final repertoires as the pairs came into breeding condition.

The videotapes were analysed to mark the occurrence of every song and to record reactions by the female during each song. Only behaviours that began after the song's onset and before its finish, an interval of less than one second, were coded. In 88 hours of recording, 18,837 songs (range per male 1,320-4,614) were recorded. On average, 94% of the songs (range 86-98%) produced no visible change in the females' behaviour. When females did respond however, a striking display was a wing stroke, that is, rapid lateral movements of one or both wings away from the body (Fig. 1). It was called to our attention in part by the alert behaviour of the males, who frequently interrupted the pace of their singing after a wing stroke to approach and inspect the subsequent behaviour of the female (Fig. 2).

Wing strokes occurred during 237 of the 18,837 songs, ranging in frequency from 2 to 70 across the eight pairs, yielding a mean of 29.6 per female and a ratio of 1.1 wing strokes to every 100

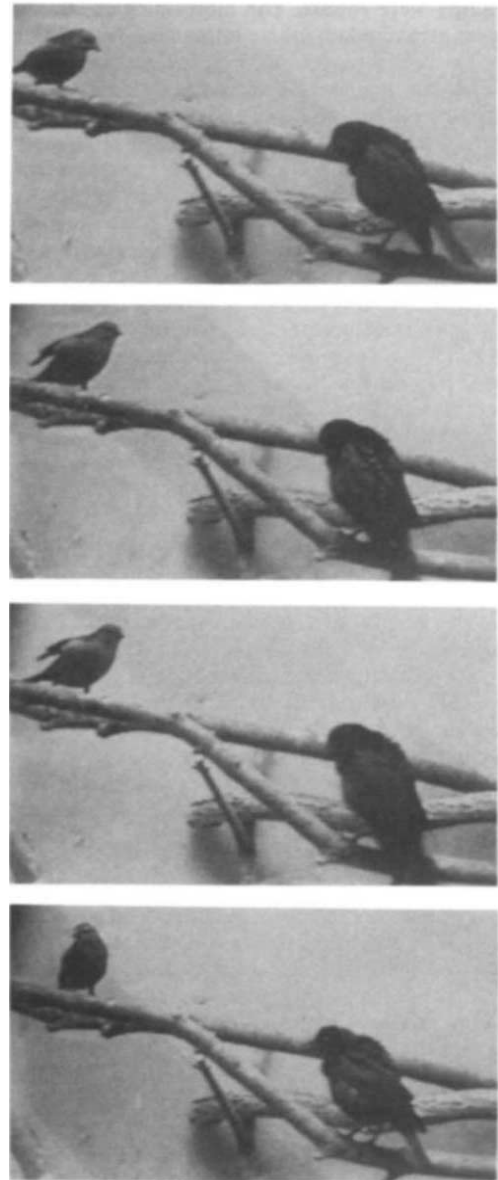


Fig. 1 A wing stroke is shown in the middle two panels. The male began his song 400 ms before the wing stroke. The upper panel shows the female's behaviour at song onset and the bottom panel shows her behaviour 900 ms later, just before the finish of the song. The photographs were made from the videotaped sequence used in experiment 4.

songs (wing strokes h^{-1} /songs h^{-1}). Two features of the wing stroke suggest it may be a precursor to a full copulatory posture, a display seen only in the breeding season. First, the wing stroke resembles the initial wing movement of a copulatory posture. Second, and most striking, the wing stroke occurs extremely rapidly while the male's one-second song is still in progress, a defining feature of the release of the copulatory posture, in this species.

The second part of the study, the playback tests, was carried out one year later. The subjects were six of the females studied in part one and two additional females which had been housed identically with males. In the time between coding of wing strokes and conclusion of playback testing, the females resided in same-sex pairs in sound-attenuating chambers.

Five song series were obtained from five of the eight videotaped male-female pairs. Each song series was composed



Fig. 2 The male's orientation and approach towards the female after producing a song that had elicited a wing stroke. The song had begun at 8:09:45:6 and ended at 8:09:46:1. The wing stroke began at 8:09:46:0 and ended at 8:09:46:2, shown in the upper panel. The female is now stretching her wing. This sequence is part of the playback series comprising experiment 3.

of 16 songs, the eight songs that had preceded a song eliciting a wing stroke, the wing-stroke song itself, and the seven songs following it. The major criterion for selecting sequences was audio recording quality. Each song series was run as a separate experiment. In each, the 16 songs were presented one at a time in six daily trials, separated in time by a minimum of 90 minutes. Each experiment took about two weeks to complete and the songs were played back an average of five times (range 4–8). The songs were not presented in their original order (hence the wing-stroke song was not necessarily the ninth song) but were ordered differently each day with the constraints that no song should occur twice in one day and that all songs should be played an equal number of times at different times of day.

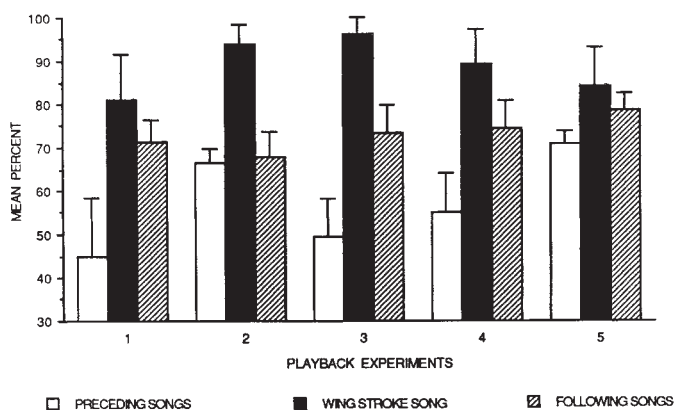


Fig. 3 The mean per cent and standard error of copulatory responses by the eight females for each experiment. Differences in potency among the 16 songs were tested by Friedman analyses of variance of ranks. Significant differences were obtained for each. The χ^2 values were as follows (df = 15): Exp. 1, 78, $P < 0.001$; Exp. 2, 31.2, $P < 0.01$; Exp. 3, 36.0, $P < 0.01$; Exp. 4, 38, $P < 0.01$; and Exp. 5, 25, $P < 0.05$. In each experiment, the wing-stroke song, song nine, received the highest mean rank. Wilcoxon signed rank tests were used to test for two differences: differences in potency between the wing-stroke song and the mean potency of songs 1–8 and differences between the mean potency of songs 1–8 and songs 10–16. For the first comparison, the Wilcoxon T value was 0, $P < 0.01$ for experiments 1–4 and $T = 2$, $P < 0.02$ for experiment 5. For the second comparison, T was 1 or 0 for experiments 1, 3, and 4, $P < 0.02$ and the T value was > 6 (n.s.), for experiments 2 and 5.

Playback testing began when the females came into breeding condition as judged by egg-laying. Song potency was defined as the percentage of trials in which a song elicited a copulatory posture, that is, the female lowered and spread her wings, arched her neck, and separated the feathers around the cloacal area.

The females responded differently to the 16 songs. In each experiment, the wing-stroke song yielded the highest mean potency, 89% (range 81–96%; Fig. 3). In contrast, the mean percent of responses to the preceding eight songs was 57% (range 0–100%) whereas the seven songs following the wing stroke elicited a mean percentage of responses of 72% (range 0–100%), producing statistically significant differences in each experiment (Fig. 3). Thus, although males sang other effective songs, only the wing-stroke song produced consistently high levels of copulatory responses.

The song sequences were examined with respect to acoustic content and organization. The five males sang four to seven different song types before the wing-stroke song. In four of the sequences (experiments 1–4, see Fig. 3), no new song types occurred after the wing-stroke song. The song type eliciting the wing stroke accounted for an average of 19% (range 0–38%) of the singing during the first eight songs in contrast to 63% (range 43–86%) of all singing after the wing stroke, a non-overlapping difference in each experiment. The song type eliciting the wing stroke was also sung more repetitively after the wing stroke: the mean number of immediate repetitions before the wing stroke was 1 (range 0–2) compared to 3.75 (range 2–6) after it. These changes, in addition to the higher potencies of songs following wing strokes, highlight the conspicuous nature of the wing stroke to males who immediately attempted to make use of the information contained in the female's display.

Thus, the present study underscores the multi-modal nature of the stimulation available to males as they learn to sing, and adds credence to a multi-phasic view of song learning whereby it is influenced by both natal and juvenile experiences⁷. Most importantly, the data bring into focus visual dimensions of song learning that have not been described before. Males must not

only listen to the nature of the sounds they produce, but also look at the reactions their sounds provoke.

Financial assistance was provided by the NSF and the National Institute of Neurological and Communicative Disorders and Strokes.

Received 11 March; accepted 31 May 1988.

1. Petrinovich, L. in *Social Learning: Psychological and Biological Perspectives* (eds Zentall, T. & Galef, B. G.) 255-278 (Lawrence Erlbaum, New York, 1988).
2. King, A. P. & West, M. J. *Nature* **305**, 704-706 (1983).
3. West, M. J. & King, A. P. *Ethology* **70**, 225-235 (1985).
4. King, A. P. & West, M. J. *Anim. Behav.* (in the press).
5. Kroodsma, D. E. in *Acoustic Communication in Birds* Vol. 2 (eds Kroodsma, D. E. & Miller, E. H.) 1-28 (Academic Press, New York, 1982).
6. West, M. J. & King, A. P. *J. comp. Psychol.* **100**, 296-303 (1986).
7. McGregor, P. K. & Krebs, J. P. *Behaviour* **79**, 127-147 (1982).

Acetylcholine inhibits identified interneurons in the cat lateral geniculate nucleus

David A. McCormick & Hans-Christian Pape

Section of Neuroanatomy, Yale University School of Medicine,
333 Cedar Street, New Haven, Connecticut 06510, USA

The transmission of visual information from retina to cortex through the dorsal lateral geniculate nucleus (LGNd) is controlled by non-retinal inputs^{1,2}. Enhanced visually evoked responses in cat LGNd relay cells during periods of increased alertness³ have been attributed in large part to increased rate of acetylcholine (ACh) release by fibres ascending from the brainstem reticular formation⁴⁻⁷. ACh can modulate geniculate visual responses *in vivo*⁵⁻⁸, but comparatively little is known about the underlying ionic mechanisms of these cholinergic actions. Although direct excitation of LGNd relay neurons has been shown *in vitro*⁹, the situation is complicated because cholinergic axons form numerous and complex synapses not only with relay cells, but also with inhibitory interneurons¹⁰, and electrical activation of the brainstem cholinergic neurons reduces inhibitory postsynaptic potentials in the LGNd¹¹⁻¹³. We report here that morphologically characterized interneurons in the cat LGNd possess distinctive electrophysiological properties in comparison with those of relay cells and are inhibited by ACh through a muscarinic receptor-mediated increase in potassium conductance. Together the direct excitation of relay cells and inhibition of intrageniculate interneurons allow the ascending cholinergic system to exert a powerful facilitatory influence over the transfer of visual information to the cerebral cortex.

Intracellular recordings from cells in lamina A or A₁ of the cat LGNd revealed two distinct classes of neurons, which we tentatively term principal (P) and I neurons. P neurons display electrophysiological properties similar to those reported previously for thalamocortical relay cells¹⁴⁻¹⁶, including the presence of a low threshold Ca²⁺ spike which underlies burst discharges (Fig. 1e, arrow head), anomalous rectification in both the hyperpolarizing (Fig. 1e) and depolarizing ranges (data not shown), a delayed onset to firing during depolarization (Fig. 1c, arrow head), and a general lack of pronounced spike frequency accommodation (Fig. 1c).

In contrast, we found that I neurons either lack completely a typical low threshold Ca²⁺ spike (Fig. 1f; *n* = 4) or display only a small rebound exultation after a large hyperpolarization (data not shown; *n* = 4). Furthermore, I neurons possess action potentials which are on the average 28% shorter in duration than those of P cells (compare Fig. 1a and b), lack a delayed onset to firing (Fig. 1d), display more linear membrane potential versus injected current (*V-I*) relationships, and have a higher input resistance (I neurons: 89 ± 24 MΩ (mean ± s.d.), *n* = 9;

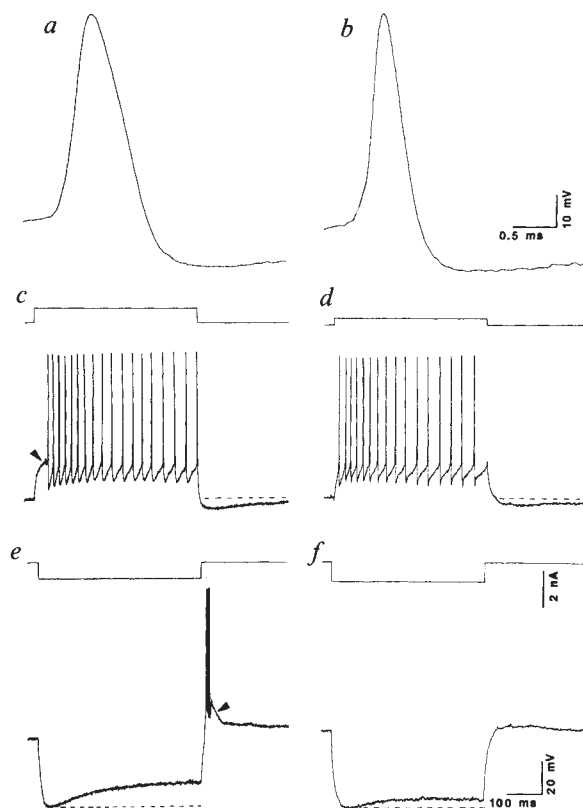


Fig. 1 Electrophysiological properties of LGNd P and I neurons. *a* and *b*, Single action potential of a typical P cell (*a*, 0.52 ms) is substantially broader at half peak amplitude than that of a typical I neuron (*b*, 0.31 ms) (average action potential duration ± s.d. for the two groups: P cells, 0.50 ± 0.20 ms, *n* = 20; I neurons, 0.36 ± 0.13 ms, *n* = 13; *t* = 2.27, *P* < 0.05). *c* and *d*, Responses to a depolarizing current pulse of a P cell (c; resting membrane potential (*V_m*) = -64 mV; arrow head indicates delayed onset of firing) and an I neuron (*d*; resting *V_m* = -76 mV). *e* and *f*, Injection of hyperpolarizing current pulses into a typical P cell (*e*) and an I neuron (*f*). Note the lack of a rebound low threshold Ca²⁺ spike in the I neuron (*f*) versus the P cell (*e*, arrow head; burst contains four action potentials) despite the large hyperpolarization from -76 to -135 mV. All data in *a*, *c* and *e* are from the same P cell and those from *b*, *d* and *f* are from the same I neuron. Calibration bars in *f* for *c-f*, calibration bars in *b* for *a* and *b*. Top trace in *c-f* is injected current, whereas the bottom trace is the resulting deviation in *V_m*.

Methods. Eleven 3-month to 3-year old cats were deeply anaesthetized with ketamine (20 mg per kg, intramuscularly) and sodium pentobarbital (15 mg per kg, intravenously) and underwent a wide craniotomy. The animals were killed by decapitation and both hemispheres were rapidly removed and dissected to isolate the lateral geniculate nucleus which was then sectioned coronally on a vibratome as 400 μm slices. Slices were maintained in an interface chamber at 35 ± 1 °C and perfused with a solution containing (in mM) NaCl, 126; KCl, 2.5; MgSO₄, 2; NaHCO₃, 26; NaH₂PO₄, 1.25; CaCl₂, 2; glucose, 10; saturated with 95% O₂, 5% CO₂ to a final pH of 7.4. Epi-illumination of the coronal slices readily revealed the main laminae of the LGNd as well as the optic tract and optic radiation. Neurons were recorded only in the middle portions of laminae A and A₁ to avoid recording from perigeniculate or interlaminar cells. Microelectrodes were filled with 4 M KAc for normal recordings and 0.2 M LiCl saturated with lucifer yellow CH (LY) for intracellular labelling experiments. Only neurons having stable membrane potentials below -55 mV and overshooting action potentials were included in the electrophysiological analysis. Action potential durations were measured at half-peak amplitude. Acetylcholine (ACh, 10 mM) or acetyl-β-methylcholine (MCh, 10 mM) were applied from a broken micropipette (1-2-μm tip diameter) by pressure pulses in volumes of approximately 10 pl within 50-75 μm of the recorded cell. LGNd slices containing LY-labelled neurons were resectioned at 50-75 μm, mounted and cleared on microscopic slides, and photographed.

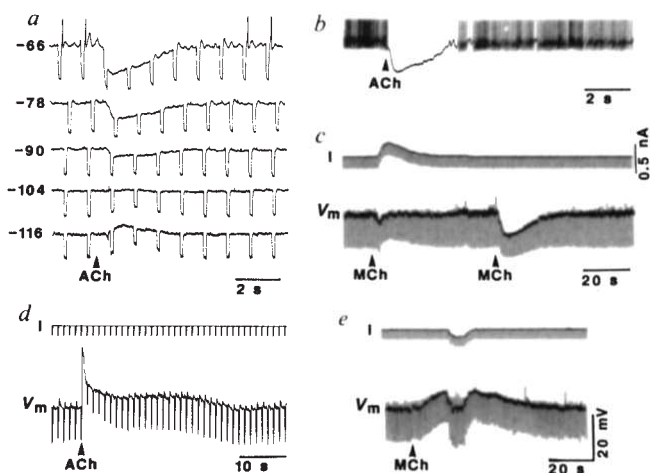


Fig. 2 Actions of acetylcholine on P and I neurons in the LGNd. *a*, Application of ACh to the I neuron illustrated in Fig. 1 results in membrane hyperpolarization which reverses to a depolarization at about -104 mV. *b*, Application of ACh to the neuron of *a* during steady depolarization to evoke spiking causes inhibition with no evidence of fast or slow excitation. *c*, MCh hyperpolarizes the I neuron (second application), and manual voltage clamp reveals the response to be associated with a substantial increase in G_i (first application). *d*, Application of ACh to a typical P cell results in a rapid depolarization followed by a slow depolarization. *e*, MCh causes the slow depolarization only (same neuron as in *d*). Compensation for the depolarization with inward current injection reveals the response to be associated with a decrease in G_i .

P cells, 51 ± 20 M Ω ; $n = 20$; $t = 4.5$, $P < 0.001$).

Application of acetylcholine to P cells at resting membrane potential causes a rapid depolarization, followed by a slower depolarization of smaller amplitude (Fig. 2*d*), as previously reported⁹. The fast depolarization, associated with a substantial increase in apparent input conductance (G_i), often elicits spike discharges, whereas the slow depolarization, associated with a decrease in G_i , reaches threshold for neuronal firing only from more depolarized membrane potentials. The muscarinic agonist, acetyl- β -methylcholine (MCh) evokes the slow depolarization and decrease in G_i only (Fig. 2*e*), which in some neurons is preceded by a small hyperpolarization⁹.

In contrast to cholinergic actions on P cells, application of ACh to electrophysiologically identified I neurons results in only a hyperpolarization and inhibition of spike discharge with no indication of fast or slow depolarizing effects ($n = 9/11$ (two cells showed no response); Figs 2*a*, *b* and 3*d*). Like ACh, the muscarinic agonist, MCh elicits a membrane hyperpolarization associated with an increase in G_i (Fig. 3*d*). Unlike reticular thalamic neurons¹⁷, the muscarinic hyperpolarization does not promote the occurrence of burst discharges in I neurons, due to an absence of a typical low threshold Ca^{2+} spike.

Possible ionic conductances underlying the ACh-induced hyperpolarization were investigated by applying ACh while the I neurons were positioned at various membrane potentials (Figs 2*a* and 3*d*). The ACh response reverses at an average membrane potential of -102.4 ± 4.8 mV ($n = 3$), which is similar to the presumed potassium equilibrium potential of thalamic neurons¹⁷. These results indicate that ACh inhibits I neurons by increasing a membrane potassium conductance mediated through muscarinic receptors, similar to that reported previously in other mammalian central nervous system neurons^{17,18}.

To test the hypothesis that the two electrophysiologically and pharmacologically defined subclasses of LGNd neurons may possess distinct morphological features, we intracellularly labelled recorded neurons with the fluorescent dye lucifer yellow CH. All of the cells demonstrating the electrophysiological

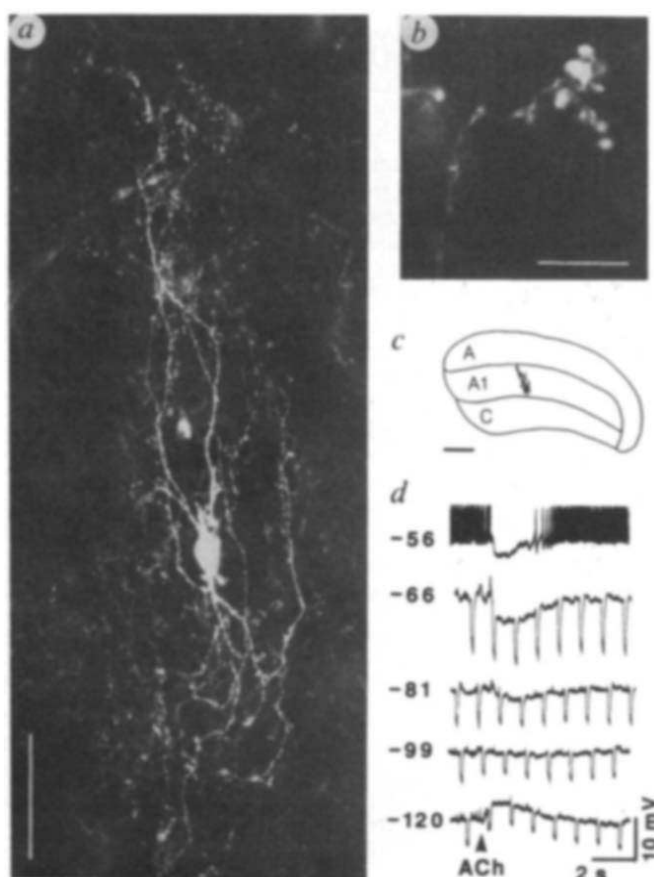


Fig. 3 Morphological characteristics of electrophysiologically and pharmacologically defined LGNd I neurons. *a*, Photomicrograph of a 75- μm section containing a LY-filled I neuron exhibiting typical type-3 morphological characteristics. Scale bar equals 50 μm . *b*, Typical terminal ending of a dendritic process of the neuron in *a*. Scale bar equals 10 μm . *c*, Position of the same neuron in the LGNd slice. Scale bar equals 300 μm . *d*, Hyperpolarizing response of the neuron in *a* to application of ACh. Reversal potential is -99 mV.

properties mentioned above for P cells were found to possess somatic and dendritic morphologies which are typical for either type-1 ($n = 14$) or type-2 ($n = 9$) relay cells as defined by Guilley¹⁹. Axons of P cells were often observed to enter into the optic radiation (data not shown). In contrast, injection of lucifer yellow into electrophysiologically identified I neurons ($n = 12$) revealed in every case cells with somatic and dendritic morphologies typical for Guilley type-3 cells^{19,20} which have recently been shown to be interneurons that contain γ -aminobutyric acid (GABA)²¹. Their morphological features included small cell bodies (type-3 cells: 17.6 ± 3.6 by 11.0 ± 3.1 μm diameter; compare with type-1 cells: 24.5 ± 3.1 by 17.3 ± 4.1 μm ; type-2 cells: 19.7 ± 3.5 by 12.4 ± 2.9 μm) and thin, sinuous, 'axon-like' dendrites which possess numerous filiform appendages ending in characteristic 'synaptic knobs' (Fig. 3*a*, *b*). None of the identified type-3 neurons possess a detectable axon which leaves the local neuronal environment. Figure 3 shows the detailed properties of a morphologically type-3 neuron. Its numerous thin dendrites arborize extensively in a bitufted manner transverse to the long axis of the LGNd (Fig. 3*a*, *c*) and possess filiform appendages and putative synaptic terminals characteristic of intrageniculate interneurons (Fig. 3*b*). Application of ACh to morphologically identified type-3 cells ($n = 5$) results in the typical hyperpolarization and associated increase in membrane conductance (Fig. 3*d*), confirming our results obtained with electrophysiologically identified I neurons.

Stimulation, microiontophoretic and neuroanatomical tracing experiments have indicated that the brainstem cholinergic system is critically involved in the ascending control of geniculate neuronal responsiveness^{1-13,22-26}. In the LGNd, acetylcholine is thought to increase relay cell excitability during periods of alertness by direct excitation and by interaction with inhibitory mechanisms^{3,5-9,22}. Our results demonstrate two distinct mechanisms by which the ascending cholinergic projection system facilitates the relay of visual information through the LGNd. The direct depolarization of relay neurons by ACh is expected to inhibit ongoing burst discharges and to bring the cell's membrane potential closer to the optimal level for the accurate transfer of incoming excitatory inputs. The inhibition of local GABAergic inhibitory interneurons by ACh will further facilitate the responsiveness of the relay neurons to excitatory inputs. Also, the inhibitory feedback loop to the LGNd via the perigeniculate nucleus, which is believed to be critical in the generation of rhythmic burst discharges of slow wave sleep²⁴, is inhibited by ACh^{5-7,17,22,23}. The presence of a widespread cholinergic innervation of the thalamus²⁵ indicates that these results may generalize to other sensory-motor systems and as such may form a basis for the ascending control of arousal.

This work was supported by the Jacob Javits Center for

Excellence in Neuroscience, and a fellowship of the Deutsche Forschungsgemeinschaft to H.-C.P.

Received 9 May; accepted 23 May 1988.

1. Singer, W. *Physiol. Rev.* **57**, 386-420 (1977).
2. Sherman, S. M. & Koch, C. *Expl. Brain Res.* **63**, 1-20 (1986).
3. Livingstone, M. S. & Hubel, D. H. *Nature* **291**, 554-561 (1981).
4. Phillips, J. W., Tebécis, A. K. & York, D. H. *J. Physiol.* **192**, 695-713 (1967).
5. Sillito, A. M., Kemp, J. A. & Berardi, N. *Brain Res.* **280**, 299-307 (1983).
6. Eysel, U. T., Pape, H.-C. & Van Schayck, R. J. *Physiol.* **370**, 233-254 (1986).
7. Francesconi, W., Müller, C. M. & Singer, W. *J. Neurophysiol.* (in the press).
8. Eysel, U. T., Pape, H.-C. & Van Schayck, R. J. *Physiol.* **388**, 199-212 (1987).
9. McCormick, D. A. & Prince, D. A. *J. Physiol.* **392**, 147-165 (1987).
10. De Lima, A. D., Montero, V. M. & Singer, W. *Expl. Brain Res.* **59**, 206-212 (1985).
11. Singer, W. & Dräger, U. *Brain Res.* **41**, 214-220 (1972).
12. Singer, W. *Brain Res.* **61**, 35-54 (1973).
13. Ahlsén, G., Lindström, S. & Lo, F.-S. *J. Physiol.* **347**, 593-609 (1984).
14. Jahnsen, H. & Llinás, R. *J. Physiol.* **349**, 205-226 (1984).
15. Jahnsen, H. & Llinás, R. *J. Physiol.* **349**, 227-247 (1984).
16. Crunelli, V., Haby, M., Jassik-Gerschenfeld, D., Leresche, N. & Pirchio, M. *J. Physiol.* **399**, 153-176 (1988).
17. McCormick, D. A. & Prince, D. A. *Nature* **319**, 402-405 (1986).
18. Egan, T. M. & North, R. A. *Nature* **319**, 405-407 (1986).
19. Guillery, R. W. *J. comp. Neurol.* **128**, 21-50 (1966).
20. Hamos, J. E., Van Horn, S. C., Raczkowski, D., Uhlrich, D. J. & Sherman, S. M. *Nature* **317**, 618-621 (1985).
21. Montero, V. M. *J. comp. Neurol.* **254**, 228-245 (1986).
22. Pape, H.-C. & Eysel, U. T. *Brain Res.* **440**, 79-86 (1988).
23. Godfraind, J. M. *Br. J. Pharmac.* **63**, 295-302 (1978).
24. Steriade, M. & Deschênes, M. *Brain Res. Rev.* **8**, 1-63 (1984).
25. Rye, D. B., Saper, C. B., Lee, H. J. & Wainer, B. H. *J. comp. Neurol.* **259**, 483-528 (1987).
26. De Lima, A. D. & Singer, W. *J. comp. Neurol.* **259**, 92-121 (1987).

Mapping of mutation causing Friedreich's ataxia to human chromosome 9

Susan Chamberlain, Jacqui Shaw, Alison Rowland, Julie Wallis, Sally South, Yusuke Nakamura*, Alexander von Gabain†, Martin Farrall & Robert Williamson

Department of Biochemistry and Molecular Genetics, Saint Mary's Hospital Medical School, University of London, London W2 1PG, UK

* Howard Hughes Medical Institute, University of Utah School of Medicine, Salt Lake City, Utah 84123, USA

† Department of Bacteriology, Karolinska Institutet, Box 60400, 104 01 Stockholm, Sweden

Friedreich's ataxia is an autosomal recessive disease with progressive degeneration of the central and peripheral nervous system^{1,2}. The biochemical abnormality underlying the disorder has not been identified. Prompted by the success in localizing the mutations causing Duchenne muscular dystrophy^{3,4}, Huntington's disease⁵ and cystic fibrosis⁶⁻⁸, we have undertaken molecular genetic linkage studies to determine the chromosomal site of the Friedreich's ataxia mutation as an initial step towards the isolation and characterization of the defective gene. We report the assignment of the gene mutation for this disorder to chromosome 9p22-CEN by genetic linkage to an anonymous DNA marker MCT112 and the interferon- β gene probe. In contrast to the clinical variation seen for the disorder, no evidence of genetic heterogeneity is observed.

The incidence of Friedreich's ataxia is approximately one in fifty thousand. Collaboration with colleagues in Europe and the United States was required to establish a collection of 22 families with three or more affected siblings, to carry out an effective linkage study. Fourteen of these pedigrees have previously been described⁹. Compliance with the diagnostic criteria of Geoffroy *et al.*¹ was observed to minimize the risk of clinical heterogeneity within the sample population.

The possibility that Friedreich's ataxia may be caused by mutations of several different genes, with one predominant mutation, has been discussed by Harding¹⁰. For the purposes

of this study we have assumed that the disease is a single-gene disorder, where variation in symptoms (such as age of onset and severity) probably result from different mutations at the same locus.

We have analysed 117 informative cloned DNA markers for linkage to Friedreich's ataxia^{9,11}, as well as 36 polymorphic blood group and protein markers, the latter in collaboration with colleagues in the United States, Denmark and Canada¹². These studies have served to exclude the disease gene from close proximity to these markers if a lod score ($\log_{10}$ odds ratio) of -2 or less was generated¹³, usually at a recombination fraction of at least $\theta = 0.10$.

Analysis of the combined data using the EXCLUDE program¹⁴, which calculates the likelihood distribution around a series of previously mapped loci, generated an exclusion map for more than 80% of the genome. Statistically, three chromosomal regions were emphasized as the most probable sites for the Friedreich's ataxia mutation. Intensive investigation of these regions led to the demonstration of linkage between the mutation causing the disease and two markers on chromosome 9, the interferon- β gene (INFB) and the anonymous probe MCT112,

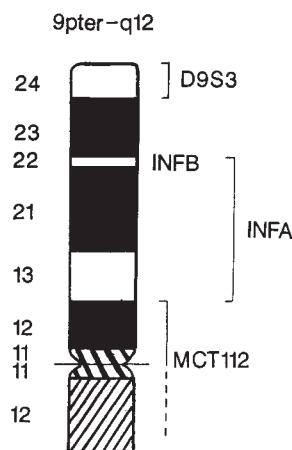


Fig. 1 Chromosome 9p22-q12. Physical assignment of the interferon- β gene and the locus defined by the DNA marker MCT112 by genetic linkage analysis.

Table 1 Pairwise lod scores between chromosome 9 markers and Friedreich's ataxia at various recombination fractions (θ), calculated using the LINKAGE computer program package²⁰

Probe	Location	Lod of θ -value						
		0.00	0.01	0.05	0.10	0.20	0.30	0.40
D9S3	pter-p24	$-\infty$	-10.64	-3.29	-0.94	0.32	0.36	0.12
INFB	9p22	$-\infty$	-0.09	2.48	2.89	2.23	1.20	0.34
MCT112	9q	6.41	6.20	5.38	4.40	2.63	1.24	0.32

both of which define MspI polymorphisms^{15,16}. At present, the INFB locus has not been genetically linked to the existing linkage map for this chromosome¹⁷. The physical locations of the markers on chromosome 9 are shown in Fig. 1.

Pairwise lod scores allowing for a male:female map distance ratio of 1:1.38 (the mean of the reported sex-differences in recombination for chromosome 9 (refs 17, 18) between Friedreich's ataxia and the DNA markers are shown in Table 1. The maximal lod score between Friedreich's ataxia and the locus defined by MCT112 calculated for combined sexes is 6.41 at a recombination fraction $\theta = 0$ (0–5% confidence interval¹⁹). No recombinants were observed between the Friedreich's ataxia locus and this probe.

Linkage was also detected with the INFB probe, with a maximal lod score of 2.98 at a male recombination fraction $\theta = 0.10$. Three-point likelihood calculations computed with the LINKAGE program²⁰ are summarized in a location map (Fig. 2). The distance between INFB and MCT112 was estimated by maximizing the joint likelihood for markers and disease. Friedreich's ataxia shows a maximal lod score of 8.99 at the locus defined by MCT112.

The interferon complex has been physically mapped to the short arm of chromosome 9 using somatic cell hybrids²¹. Studies on patients with acute monocytic leukaemia and a balanced translocation t(9; 11)(p22; q23)²² show that the breakpoint on 9p splits the interferon genes and more precisely localizes the interferon- β gene to 9p22 and the interferon- α gene to 9p22-p12. The anonymous probe MCT112 has been assigned to proximal chromosome 9q by multipoint linkage analysis with loci (D9S1, ASSP3 and ABO)^{17,23}, but has not been physically localized.

We have shown that INFB, MCT112 and Friedreich's ataxia comprise a linkage group spanning ~10 centimorgans (cM), although this may represent a large proportion of chromosome 9 due to suppression of recombination around the centromere. The physical locations for D9S1 and ASSP3 are 9pter-q11 and q11-q22 respectively, and MCT112 is located by linkage at a

distance of 12–15 cM from these loci towards 9pter. We tentatively assign MCT112 and Friedreich's ataxia to 9p22-centromere. This is supported by the exclusion data generated between the Friedreich's ataxia locus and distal chromosome 9q markers previously reported¹¹.

The linkage data, and the inclusion in the study of families from several geographically isolated populations, argues strongly in favour of a single gene locus for the classical form of the disorder. No evidence for heterogeneity was detected, with most families being informative for at least one of the linked loci. The incidence of cardiomyopathy in these families remains to be studied in detail, but where data are available this pathology is observed. Using the linked markers it should be possible to re-assess the diagnosis in cases of 'atypical' Friedreich's ataxia, which do not exhibit all of the classical symptoms of the disorder, to determine whether or not they are caused by a mutation at the same locus.

Extension of the linkage map to give flanking markers is a high priority, as is the orientation of the map with respect to existing markers on chromosome 9p. Pre-symptomatic diagnosis, carrier detection and pre-natal diagnosis can then be offered to those families with affected individuals.

Several large inbred families exist which may permit the technique of homozygosity mapping²⁴ to be applied in conjunction with other existing strategies to move from linkage to gene. This assumes that the region surrounding the disease locus remains homozygous by descent in affected children of consanguineous marriages. A group of related Friedreich's ataxia families in Cyprus was recently described by Dean *et al.*²⁵; as these families are expected to be descended from a single common ancestor, they should provide an excellent resource to study convergence of homozygosity.

We thank the Friedreich's Ataxia Groups of the United Kingdom and Eire, and the families for support and encouragement, Drs A. Harding, M. McLaughlin, S. Hawkins, S. di Donato, M. Filla, M. Anvret, T. Deufel, K. Fischbeck and C. Coutelle for help with the collection of the families used in this study, and the groups which donated the probes not showing linkage. Mansoor Sarfarazi helped us analyse the data using the EXCLUDE program. Finally we thank Bronya Keats, Keith Johnson and Peter Jones for support and discussions. This project is funded by the Friedreich's Ataxia Group, UK, the Wolfson Trust and the Muscular Dystrophy Association of America.

Received 26 April; accepted 14 June 1988.

- Geoffroy, G. *et al.* *Can. J. Neurol. Sci.* **3**, 279–287 (1976).
- Hewer, R. L. *Br. med. J.* **3**, 649–652 (1968).
- Davies, K. E. *et al.* *Nucleic Acids Res.* **8**, 2303–2312 (1983).
- Monaco, A. P. *et al.* *Nature* **323**, 646–650 (1986).
- Gusella, J. F. *et al.* *Nature* **306**, 234–238 (1983).
- Tsui, L.-C. *et al.* *Science* **230**, 1054–1057 (1985).
- Wainwright, B. J. *et al.* *Nature* **318**, 384–385 (1985).
- White, R. *et al.* *Nature* **318**, 382–384 (1985).
- Chamberlain, S. *et al.* *Hum. Genet.* **76**, 186–190 (1987).
- Harding, A. E. *Brain* **104**, 589–620 (1981).
- Chamberlain, S., Shaw, J., South, S., Farrall, M. & Williamson, R. *Cytogenet. Cell Genet.* **46**, 591 (1988).
- Keats, B. J. B. *et al.* *Am. J. hum. Genet.* **41**, 627–634 (1987).
- Morton, N. E. *Am. J. hum. Genet.* **7**, 277–318 (1955).
- Sarfarazi, M., Huson, S. & Edwards, J. E. *J. med. Genet.* **24**, 515–520 (1987).
- Ohlsson, M., Feder, J., Cavalli-Sforza, L. L. & von Gabain, A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4473–4476 (1985).
- Carlson, M. *et al.* *Nucleic Acids Res.* **15**, 10614 (1987).

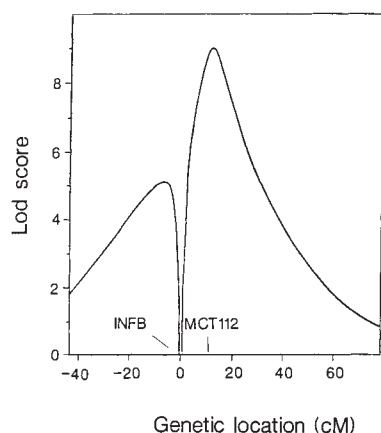


Fig. 2 Location map summarizing lod scores ($\log_{10}$ odds ratio) calculated for Friedreich's ataxia at various map positions in a fixed marker map. The relative genetic position of INFB arbitrarily has been put at 0.

17. Lathrop, G. M. *et al. Cytogenet. Cell Genet.* **46**, 644 (1988).
18. Donis-Keller, H. *et al. Cell* **51**, 319-337 (1987).
19. Conneally, P. M. *et al. Cytogenet. Cell Genet.* **40**, 1-4, 356-359 (1985).
20. Lathrop, G. M., Lalouel, J. M., Julien, C. & Ott, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3443-3446 (1984).
21. Owerback, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3123-3127 (1981).
22. Diaz, M. O., Le Beau, M. M., Pitha, P. & Rowley, J. D. *Science* **231**, 265-267 (1986).
23. Smith, M. & Spence, A. *Cytogenet. Cell Genet.* **40**, 156-178 (1985).
24. Lander, E. S. & Botstein, D. *Science* **236**, 1567-1570 (1987).
25. Dean, G., Chamberlain, S. & Middleton, L. *Lancet* **i**, 587 (1988).

NMDA application potentiates synaptic transmission in the hippocampus

Julie A. Kauer, Robert C. Malenka & Roger A. Nicoll*

Departments of Pharmacology & Physiology,
University of California, San Francisco, California 94143, USA

The NMDA (*N*-methyl-D-aspartate) class of glutamate receptor plays a critical role in a variety of forms of synaptic plasticity in the vertebrate central nervous system¹⁻⁵. One extensively studied example of plasticity is long-term potentiation (LTP), a remarkably long-lasting enhancement of synaptic efficiency induced in the hippocampus by brief, high-frequency stimulation of excitatory synapses. LTP is a strong candidate for a cellular mechanism of learning and memory. The site of LTP induction appears to be the postsynaptic cell and induction requires both activation of NMDA receptors by synaptically released glutamate⁶ and depolarization of the postsynaptic membrane⁷⁻⁹. It is proposed that this depolarization relieves a voltage-dependent Mg²⁺ block of the NMDA receptor channel, resulting in increased calcium influx^{10,11} which is the trigger for the induction of LTP^{1,12,13}. This model predicts that application of a large depolarizing dose of NMDA should be sufficient to evoke LTP. In agreement with a previous study⁶, we have found that NMDA or glutamate application does potentiate synaptic transmission in the hippocampus. This agonist-induced potentiation is, however, decremental and short-lived, unlike LTP. It is occluded shortly after the induction of LTP and a similar short-term potentiation can be evoked by synaptically released glutamate. We thus propose that LTP has two components, a short-term, decremental component which can be mimicked by NMDA receptor activation, and a long-lasting, non-decremental component which, in addition to requiring activation of NMDA receptors, requires stimulation of presynaptic afferents.

Iontophoresis of NMDA in hippocampal slices routinely produced a transient depression of the excitatory postsynaptic potential (e.p.s.p.) followed by a marked potentiation (Fig. 1, maximum increase in slope of the e.p.s.p. = 63.3 ± 6.8%, *n* = 40). Presynaptic fibre stimulation in the presence of the agonist, was not required for the potentiation and NMDA could elicit potentiation repeatedly in the same slice if sufficient time (30-45 min) was allowed between applications. A potentiation of similar magnitude was observed using intracellular recording (*n* = 15) and this potentiation was not associated with any change in membrane potential or input resistance. In contrast, as previously reported⁶, iontophoresis of quisqualic acid, which selectively activates a separate class of glutamate receptors on the pyramidal cells^{14,15}, produced a transient depression but failed to cause any potentiation (*n* = 6). Furthermore, iontophoresis of quisqualic acid and NMDA together produced an increase in the e.p.s.p. that was indistinguishable from that produced by NMDA alone (*n* = 6). Iontophoresis of glutamate produced a transient enhancement of the e.p.s.p. (maximum increase in slope of the e.p.s.p. = 47.0 ± 4.1%, *n* = 12) similar to that elicited by NMDA. These data suggest that agonist-induced

potentiation results from the specific activation of NMDA receptors and that activation even of the full complement of glutamate receptors fails to produce a long-lasting non-decremental form of potentiation. The evidence favours a postsynaptic site for the potentiating effects of NMDA¹⁶, although presynaptic effects cannot be entirely excluded.

On average, the agonist-induced potentiation gradually decayed back to baseline values within 10-30 min. In contrast, a high-frequency tetanus produced a potentiation which remained at values considerably above baseline for well over an hour. Figure 1*a* shows that, following recovery to control values after agonist-induced potentiation, a tetanus can still elicit LTP with a clear long-lasting non-decremental component not evident in the NMDA-induced potentiation. In addition, in many cases a decremental potentiation was observed superimposed on the long-lasting component. To compare in more detail agonist-induced potentiation and tetanus-induced LTP we measured the magnitude of synaptic enhancement at several time points. As shown in Fig. 1*b*, the difference between the two forms of potentiation became apparent by 30 min.

Are the processes which underlie agonist-induced potentiation relevant to LTP? We predicted that if the mechanisms responsible for agonist-induced potentiation were also activated by glutamate released during tetanus-induced LTP, the potentiating effect of the agonist should be occluded shortly after tetanic stimulation. Figure 2 demonstrates that this was, in fact, observed. Initially, iontophoresis of NMDA caused a typical short-lasting potentiation. Shortly after maximal activation of the processes underlying LTP by multiple tetani, NMDA application failed to evoke a potentiation. In contrast, 60 min after the tetani, at a time when the e.p.s.p. remained strongly potentiated, re-application of NMDA once again elicited a potentiation identical to that observed with the first exposure to NMDA (*n* = 4).

If the agonist-induced potentiation is of physiological relevance, it should be possible to duplicate the above experiments by activating NMDA receptors with synaptically released glutamate. To test this possibility, we examined the effects of LTP induced by multiple tetani on the responses to a subsequent tetanus. Figure 3 shows the results of a typical experiment. Shortly after the repetitive tetani, a further tetanus failed to evoke any potentiation other than post-tetanic potentiation. In contrast, 60 min after the tetani, at a time when the e.p.s.p. remained strongly potentiated, a subsequent tetanus elicited a short-lasting decremental potentiation of the e.p.s.p. with a time course similar to that observed with NMDA application (*n* = 4). The results of the preceding two experiments suggest that the mechanisms which underlie agonist-induced potentiation may be activated by synaptically released glutamate during the induction of LTP. Furthermore, these mechanisms are not activated for the duration of LTP, but decay after LTP induction and, if sufficient time has elapsed, can be reactivated in synapses still exhibiting LTP.

Recent studies have suggested that protein kinase C plays an important role in LTP¹⁷⁻²⁰. In particular, in the presence of H-7 or sphingosine, inhibitors of this enzyme, the potentiation following a tetanus is decremental, returning to control levels within an hour^{21,22}. The similar time course of tetanus-induced potentiation elicited in the presence of protein kinase C inhibitors, and agonist-induced potentiation, suggests that they may result from the same protein kinase C-independent mechanism. As shown in Fig. 4*a*, agonist-induced potentiation was little affected by a concentration of sphingosine sufficient to block LTP (Fig. 4*b*) (see refs 22, 23). Results identical to those in Fig. 4*a* were obtained using a structurally different protein kinase C antagonist, H-7, (300 μM) (*n* = 5). Thus, protein kinase C is unlikely to be required for the potentiation elicited by NMDA receptor activation.

Previous studies with the NMDA antagonist D-2-amino-5-phosphonovalerate (AP5) have shown that NMDA receptor

* To whom correspondence should be addressed.

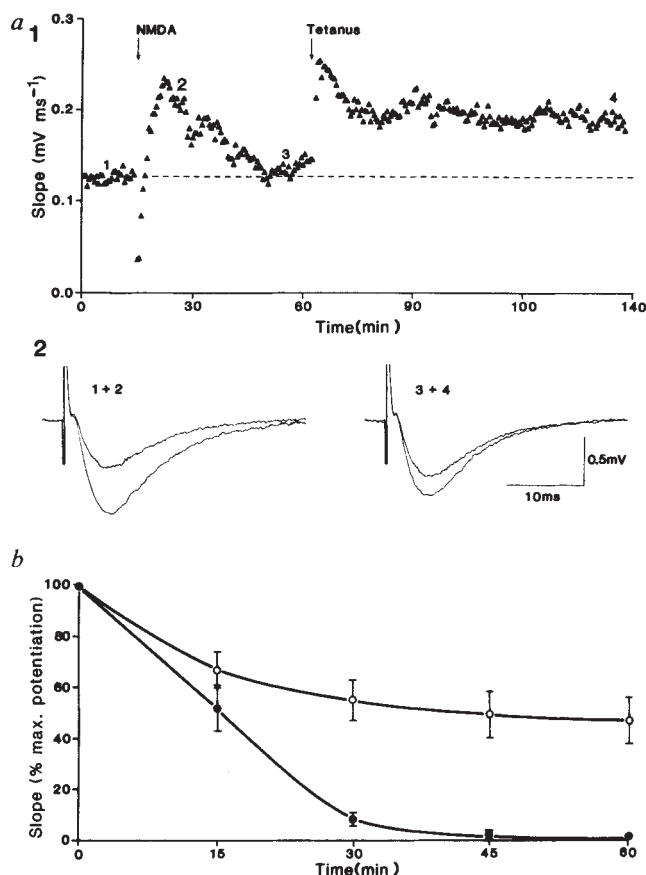


Fig. 1 Agonist-induced potentiation differs from tetanus-induced LTP. *a*₁, The slope of the e.p.s.p. is plotted as a function of time. At the time indicated by the first arrow, stimulation was halted and NMDA was iontophored (10 s, 400 nA). Stimulation was resumed 1 min after the application of NMDA. After the e.p.s.p. returned to baseline values, the afferent pathway was tetanically stimulated (second arrow). Each point on this and subsequent graphs represents the average of three slope measurements. *a*₂, Field e.p.s.p. records were taken at the times indicated by the numbers in (*a*₁) (records are the average of six sweeps). *b*, Relative potentiation observed following iontophoresis of NMDA or tetanic stimulation. Maximal potentiation following a tetanus was defined as the potentiation measured 5 min after the tetanus, to exclude contamination by post-tetanic potentiation. The maximum potentiation as a percentage of baseline was $63.3 \pm 6.8\%$ ($n=40$) following NMDA application, and $90.6 \pm 7.6\%$ ($n=23$) following tetanic stimulation. The potentiation remaining at each time point is expressed as a percentage of maximum potentiation. Data are expressed as the mean $\pm$ s.e.m. Significance was assessed using Student's *t*-test. 15 min: NMDA, $n=40$; tetanus, $n=21$; $P<0.05$; 30 min: NMDA, $n=38$; tetanus, $n=22$; $P<0.0005$; 45 min: NMDA, $n=38$; tetanus, $n=22$; $P<0.0005$; 60 min: NMDA, $n=38$; tetanus, $n=20$; $P<0.0005$. All tetanized slices were included for the purposes of this comparison, even though some slices failed to exhibit potentiation at 60 min after tetanus.

Methods. Hippocampal slices were prepared as previously described²⁹. All experiments involved stimulation of the Schaffer collateral/commissural afferents at 0.1 Hz during recording of the field e.p.s.p. (200–500 μ V original amplitude) in stratum radiatum in the CA1 region. NMDA (100 mM, pH 8), glutamate (1 M, pH 8), or quisqualate (50 mM, pH 8) was applied iontophoretically (150–400 nA for 5–20 s) from one barrel of a double-barrelled extracellular recording electrode, to ensure that the recorded field potentials originated within the synaptic region receiving NMDA. Tetanic stimulation was delivered at 1.5 or 2.0 times the intensity of the test pulse, for 1 s at 100 Hz, two times, 20 s apart. Data were collected and analysed using a modified version of pClamp (Axon Instruments) which measured the initial slope of the field e.p.s.p. Experiments were carried out either in standard solution¹⁷ or in picrotoxin-containing solution (in mM: NaCl, 124; KCl, 4; CaCl₂, 4; MgCl₂, 4; NaCO₃, 26; glucose, 10; picrotoxin, 0.1)⁹. When the picrotoxin-containing solution was used, a surgical cut was made to isolate the CA3 region from CA1, to limit epileptiform discharges. Results using each solution are pooled as no obvious differences were observed.

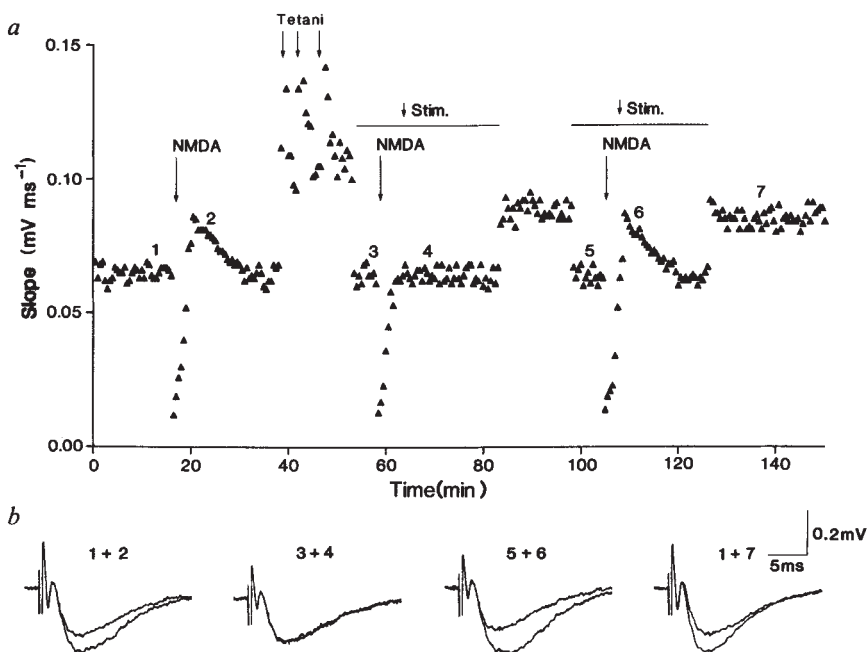


Fig. 2 NMDA-induced potentiation is occluded shortly after the induction of LTP, but can be elicited again within 60 min of the tetanus. *a*, NMDA (200 nA for 5 s) was iontophored at the times shown. Three successive tetani activated LTP maximally. Immediately following the tetanus, and at 50 min post-tetanus, the stimulus strength was reduced to produce an e.p.s.p. comparable to control levels (times are denoted with bars) and NMDA was again iontophored. *b*, Representative field e.p.s.p. records were taken at the times indicated by numbers in (*a*). Records are the averages of six sweeps.

activation is necessary for LTP^{1,14,24}. The present studies with NMDA application indicate that, although NMDA receptor activation does cause a short-term potentiation, it is not sufficient for the induction of LTP. Furthermore, our results suggest that tetanic stimulation causes two forms of NMDA-dependent potentiation: first, a short-term, decremental component which

resembles the potentiation produced by NMDA application; and second, a non-decremental component which has been shown previously to be blocked by AP5, indicating that it also requires NMDA receptor activation⁶. An issue which remains to be resolved is whether the short form of potentiation caused by NMDA receptor activation is an essential step in the estab-

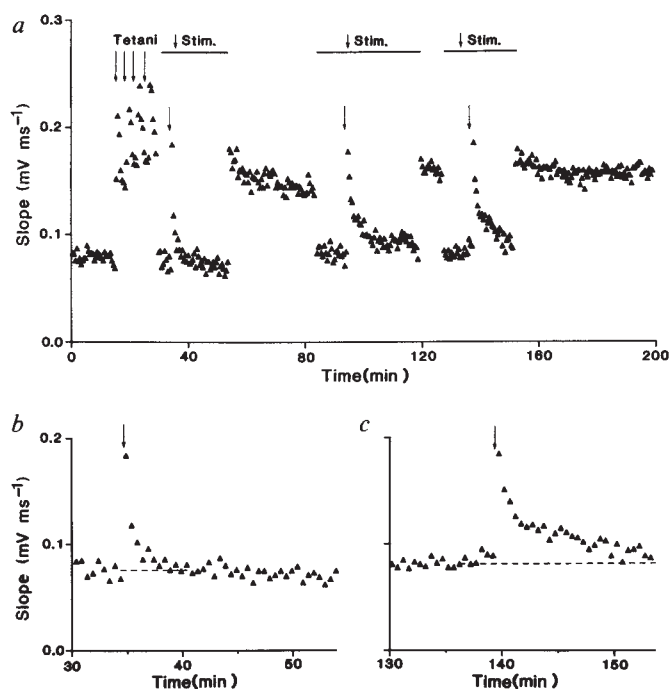


Fig. 3 A short-term, tetanus-induced potentiation is occluded shortly after the induction of LTP, but can be observed 60 and 100 min later. Four successive tetani were used to activate LTP maximally. Immediately after the tetanus and at 60 and 100 min later, the stimulus strength was reduced to elicit an e.p.s.p. comparable in size to those at the baseline (at times denoted by bars), and a further tetanus was delivered. *b* and *c*, Expanded timescale showing the potentiation of the e.p.s.p. elicited by a tetanus immediately following LTP induction (*b*) and a tetanus 100 min after LTP induction (*c*).

lishment of LTP or is a parallel event. In either case, this short-term form of information storage may be used under physiological conditions.

The present results indicate that presynaptic stimulation provides an essential ingredient for LTP. A number of possibilities exist. First, activity in the presynaptic terminal may render the synapses receptive to some factor, perhaps released from the postsynaptic neuron²⁴. Second, an additional factor, either co-released from the glutamate-containing synaptic terminals or released from a separate population of afferents, is required. Norepinephrine has been proposed to serve such a function in the generation of LTP at mossy fibre synapses in hippocampal region CA3²⁵, although this form of LTP does not require NMDA receptor activation^{26,27}. In the striate cortex, the development of ocular dominance can be prevented by AP5².

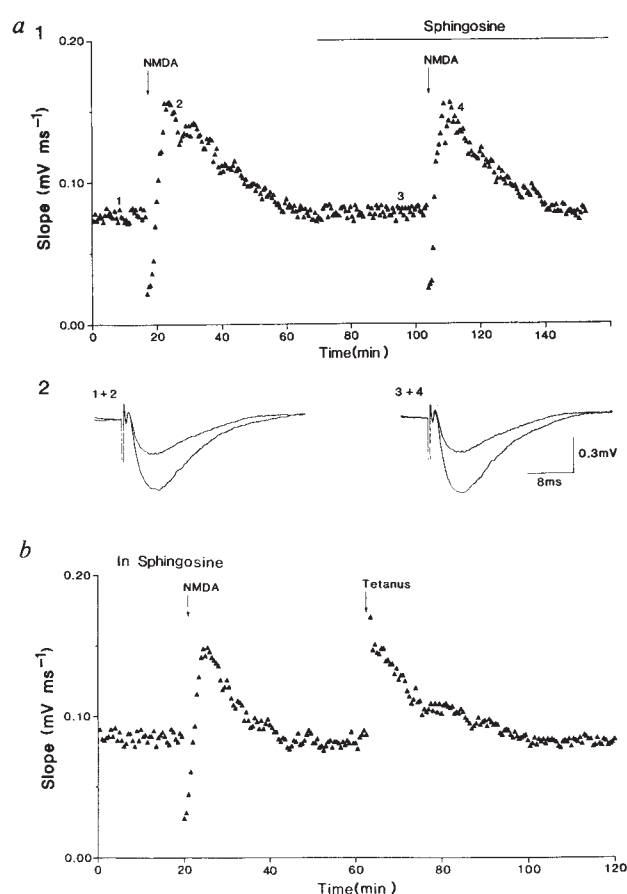


Fig. 4 Inhibition of protein kinase C does not block agonist-induced potentiation. *a*₁, The e.p.s.p. potentiation elicited by NMDA application before and 35 min after superfusing the slice with a solution containing sphingosine (10 μM). *a*₂, Sample records obtained at the times indicated by number in *a*₁ (records are the average of six sweeps). *b*, Experiment carried out in sphingosine (10 μM) in which both NMDA application and tetanic stimulation elicit a decremental potentiation. The sphingosine solution was made from a 10 mM stock in ethanol. The solution was sonicated for 5 min before application.

This form of plasticity also appears to require additional transmitters²⁸. It will be of interest to determine whether additional factors are required by other forms of NMDA-dependent synaptic plasticity. Based on the sensitivity of LTP to protein kinase C inhibitors²¹⁻²³, one might predict that such factors would activate protein kinase C¹⁷.

We thank D. Perkel for his help in setting up the computer routines for storage and analysis of the data and D. Perkel and Dr P. Sah for comments on the manuscript. This work was supported by grants from the NIH.

Received 5 April; accepted 24 May 1988.

- Collingridge, G. L. & Bliss, T. V. P. *Trends Neurosci.* **10**, 288-293 (1987).
- Kleinschmidt, A., Bear, M. F. & Singer, W. *Science* **238**, 355-358 (1987).
- Cline, H. T., Debski, E. A. & Constantine-Paton, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4342-4345 (1987).
- Stelzer, A., Slater, N. T. & tenBruggencate, G. *Nature* **326**, 698-701 (1987).
- Mody, I. & Heinemann, U. *Nature* **326**, 701-704 (1987).
- Collingridge, G. L., Kehl, S. J. & McLennan, H. J. *Physiol. Lond.* **334**, 33-46 (1983).
- Gustafsson, B., Wigström, H., Abraham, W. C. & Huang, Y.-Y. *J. Neurosci.* **7**, 774-780 (1987).
- Malinow, R. & Miller, J. P. *Nature* **320**, 529-530 (1986).
- Gustafsson, B. & Wigström, H. *J. Neurosci.* **6**, 1575-1582 (1986).
- Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 262-263 (1984).
- Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462-465 (1984).
- Lynch, G., Larson, J., Kelso, S., Barriouneuve, G. & Schottler, F. *Nature* **305**, 719-721 (1983).
- Wigström, H. & Gustafsson, B. *J. Physiol., Paris* **81**, 228-236 (1986).

- Collingridge, G. L., Kehl, S. J. & McLennan, H. J. *Physiol. Lond.* **334**, 19-31 (1983).
- Watkins, J. C. & Evans, R. H. *A. Rev. Pharmac. Tox.* **21**, 165-204 (1981).
- Kauer, J. A., Malenka, R. C. & Nicoll, R. A. *J. Physiol., Lond.* **398**, 23P (1988).
- Malenka, R. C., Madison, D. V. & Nicoll, R. A. *Nature* **321**, 175-177 (1986).
- Akers, R. E., Lovinger, D. M., Colley, P. A., Linden, D. J. & Routtenberg, A. *Science* **231**, 587-489 (1986).
- Malenka, R. C., Ayoub, G. S. & Nicoll, R. A. *Brain Res.* **403**, 198-203 (1987).
- Hu, G.-Y. *et al. Nature* **328**, 426-429 (1987).
- Lovinger, D., Wong, K., Murakami, K. & Routtenberg, A. *Brain Res.* **436**, 177-183 (1987).
- Malinow, R., Madison, D. V. & Tsien, R. W. *Biophys. J.* **53**, 429a (1988).
- Madison, D. V., Malinow, R. & Tsien, R. W. *J. Physiol., Lond.* **398**, 18P (1988).
- Bliss, T. V. P. & Lynch, M. A. in *Long-term Potentiation: Mechanisms and Key Issues* (eds Landfield, P. W. & Deadwyler, S. A.) 3-72 (Liss, New York, 1988).
- Hopkins, W. F. & Johnston, D. *Science* **226**, 350-352 (1984).
- Harris, E. W. & Cotman, C. W. *Neurosci. Lett.* **70**, 132-137 (1986).
- Kauer, J. A. & Nicoll, R. A. *Exp. Brain Res.* (suppl.) (in the press).
- Bear, M. F. & Singer, W. *Nature* **320**, 172-175 (1985).
- Nicoll, R. A. & Alger, B. E. *J. Neurosci. Meth.* **4**, 153-156 (1981).

Can B cells turn on virgin T cells?

Olli Lassila, Olli Vainio & Polly Matzinger

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland

The first event in the initiation of an immune response is the capture and presentation of antigen to T cells. Such presentation involves two distinct steps: (1) display of the antigen, which requires uptake, processing and re-expression of the antigen in association with MHC molecules on the presenting cell surface; and (2) triggering, in which the presenting cell provides signals leading to the activation of the responding T cell. Two sorts of cells can capture antigens, the 'professional' antigen-presenting cells (APCs) such as dendritic cells and macrophages, and the B cells. Both types of cells can display antigens and the APCs are known to be able to trigger resting T cells. But despite *in vitro* evidence that certain B-cell types can reactivate previously-activated T cells¹⁻⁶, it is not yet clear whether a B cell can initiate an immune response by providing the signals necessary to activate a resting T cell. We reasoned that resting B cells should not have

this capacity because of the problems this would present with tolerance to self idiotypes. By exploiting the unique properties of the avian haematopoietic system, we have examined the presenting capacity of B cells *in vivo* and found that resting B cells are indeed unable to activate resting T cells.

It is possible in chickens, as it is not in mammals, to construct an A→B chimera containing T cells and APCs of the host 'B' type strain and B cells from an allogeneic donor 'A' strain^{7,8}. Chickens treated neonatally with cyclophosphamide will regenerate their own T cells and APCs but not B cells⁷⁻⁹ because they no longer have stem cells capable of homing to the bursa. However, an injection of committed B cell precursors from normal bursae restores their ability to generate a mature B cell compartment^{8,9}. Such B-cell chimaeric chickens are tolerant of the donor alloantigens. They do not reject donor type skin grafts^{7,10}, nor do they respond in graft vs host assays or mixed leukocyte cultures¹⁰. But when the host and donor strains differ in the class II region of the major histocompatibility complex (MHC) the chimaeras cannot mount T-dependent antibody responses (refs 7, 8 and Table 1).

Table 1 shows the responses of cyclophosphamide-treated chickens which have been restored with syngeneic bursal stem cells (B21→B21 and B15→B15), allogeneic bursal stem cells differing at all known chicken MHC loci (B21→B15 and B15→

Table 1 B-cell responses in syngeneic and allogeneic chicken chimaeras

Donor→host	Anti <i>Brucella</i> agglutination titre	Anti-SRBC haemagglutination titre		Anti-SRBC PFCs/ 10 ⁶ spleen cells		Number of germinal centres	Protein A PFCs/ 10 ⁶ spleen cells
		Total	IgG	Total	IgG		
B21→B21	1,024	4,096	64	2,560	1,950	46	11,350
	4,096	8,192	128	4,380	2,640	52	18,640
	1,024	1,024	16	1,200	320	17	7,890
	512	512	16	2,960	870	41	10,520
	256	512	32	3,040	740	22	8,230
	1,024	4,096	128	3,674	3,140	49	14,770
B15→B15	4,096	1,024	64	4,550	3,170	48	18,360
	1,024	256	16	1,280	820	13	7,880
	8,192	2,048	128	5,990	3,740	45	15,290
	1,024	512	32	1,890	760	31	9,420
	1,024	128	32	2,180	1,630	23	10,430
B21→B15	512	4	0	40	0	1	8,230
	256	2	0	0	0	3	5,190
	1,024	2	0	40	0	0	9,930
	1,024	4	0	240	0	2	10,790
	512	4	0	80	0	1	4,130
	256	2	0	20	0	0	7,730
B15→B21	1,024	4	0	0	0	2	10,710
	2,048	8	0	40	0	0	8,250
	128	2	0	0	0	0	4,930
	2,048	8	0	40	0	3	12,740
	256	4	0	20	0	2	6,250
(—)→B21	0	0	0	0	0	0	150
	2	4	0	0	0	2	220
	0	0	0	0	0	0	0
	0	0	0	0	0	2	80
(—)→B15	0	0	0	0	0	1	160
	0	2	0	0	0	0	120
	0	0	0	0	0	0	80
	0	0	0	0	0	1	40

The H.B21 (B21) and H.B15 (B15) chicken strains differ in all known MHC loci¹¹ and have been described elsewhere¹². Chickens were maintained at Gifp Oberfrick, Switzerland. To make chimaeras, recipient hosts received 3.0 mg (B15) or 4.0 mg (B21) cyclophosphamide (Läake OY, Turku, Finland) intraperitoneally (i.p.) daily for the first four days after hatching. On the fourth day they were injected intravenously with 2×10^7 normal four-day-old bursa cells, prepared as described^{7,8}. Five and six weeks after hatching the chickens were injected i.p. with 10^8 sheep red blood cells (SRBC) and 10^7 *Brucella abortus* organisms. Their immune responses were measured five days after the second injection. Antibody titres were assayed by haemagglutination of SRBC and *Brucella* agglutination⁷. To measure IgG responses, 25 μ l of serum was incubated with 25 μ l of 0.2 M β -mercaptoethanol for 1 h at 37 °C before the haemagglutination assay. Anti-SRBC plaque-forming cells (PFCs) were counted⁸ using rabbit anti-chicken immunoglobulin (R α ChIg) and goat anti-chicken IgG (Nordic, Tilburg, the Netherlands) to develop total or IgG plaques respectively. Protein A PFCs were enumerated using R α ChIg as developing antibody¹³. The number of PFCs is expressed as mean value of triplicate assays. We counted the number of germinal centres in one cross section at the largest circumference of the spleen, as described elsewhere⁷.

Table 2 B-cell responses following injection of donor or host type APCs

Donor → Host		MHC type of injected APCs	Anti-SRBC haemagglutination titre		Anti-SRBC PFCs/ 10 ⁶ spleen cells		Number of germinal centres	Protein A PFCs/ 10 ⁶ spleen cells
			Total	IgG	Total	IgG		
<i>a, Injected with donor type APCs</i>								
B15 → B21		B15	4,096	16	10,850	3,290	12	13,200
		B15	64	8	6,060	4,920	26	27,040
		B15	256	8	6,100	4,960	32	47,040
		B15	256	256	3,160	2,840	17	14,080
B21 → B15		B21	2,048	16	3,040	1,870	16	17,360
		B21	1,024	512	5,090	3,040	12	7,360
<i>b, Injected with host type APCs</i>								
B15 → B21		B21	2	0	20	0	0	4,560
		B21	4	0	20	0	1	7,540
		B21	4	0	30	0	2	17,200
		B21	0	0	ND	ND	ND	ND
B21 → B15		B15	8	0	230	0	1	8,640
		B15	2	0	20	0	2	5,960

Chimaeras were immunized and tested as for Table 1. Those which had restored their B-cell compartments, as evidenced by their ability to respond to the T-independent antigen, *Brucella*, were used at the age of eight weeks. They were immunized again with 10⁵ SRBC mixed together with 5–7 × 10⁸ spleen cells from unimmunized, age-matched chickens. The spleen cells were irradiated (3,000 rad), mixed with SRBC in Dulbecco's minimal Eagles medium and left at room temperature for 2–4 h before injection. Two injections were given seven days apart and five days later the immune responses were analysed as for Table 1. ND, not determined.

Table 3 B-cell responses following injection of various types of APCs

Donor → Host	MHC type of injected APCs	Anti SRBC haemagglutination titre			
		Primary response		Secondary response	
		Total	IgG	Total	IgG
i) Injected with donor APCs					
B21 → B15	B21	64	0	8,192	16
B21 → B15	B21	8	0	512	8
B15 → B21	B15	16	0	8,192	8
ii) (Donor × host)F₁ APCs					
B15 → B21	B15 × B21	16	0	8,192	8
B15 → B21	B15 × B21	64	8	512	8
B15 → B21	B15 × B21	64	0	8,192	32
iii) (Donor × third party)F₁ APCs					
B21 → B15	B12 × B21	64	8	256	64
B21 → B15	B12 × B21	0	0	256	0
B15 → B21	B12 × B15	32	0	1,024	32
iv) (Host × third party)F₁ APCs					
B21 → B15	B12 × B15	0	0	0	0
B15 → B21	B12 × B21	0	0	4	0
v) Third party APCs					
B21 → B15	B12	0	0	0	0
B15 → B21	B14	8	0	4	0

The construction of bursa-cell chimaeras was as described in Table 1. The immunization was carried out as described in Table 2 using spleen cells from different inbred strains of chickens. The nomenclature and strains used have been described¹². B12 and B14 haplotype cells were from the CB and H.B14 strains. Serum antibody measurements were as described in Table 1. The immunization started at the age of 6–9 weeks after hatching. Primary response was determined seven days after first immunization and secondary response five days after second immunization.

B21) or nothing. It is clear from the responses to the T-independent antigen, *Brucella abortus*, that A → B allogeneically reconstituted chickens have functional B cells, yet their T-dependent responses to sheep red blood cells (SRBC) are hardly distinguishable from those of the totally unreconstituted birds, whether expressed as serum titres, plaque forming cells (PFC) or numbers of germinal centres. Clearly there is a breakdown in T-B cell communication in the allogeneic A → B chimaeras.

There are two potential reasons for the defective response. (1) Thymic education: it is generally, though not universally^{14,15}, believed that the T-cell repertoire is positively selected by the MHC antigens of the thymus such that T cells which have matured in an MHC type A thymus will be able to recognize antigens only in association with MHC 'A' and not with MHC

'B'^{16,17}. Thus the T cells in the A → B chimaeras, having matured in the host 'B' type thymus would be unable to collaborate with the 'A' type B cells. But because thymic education is not absolute^{18–21}, it is possibly not the principal defect. (2) Lack of appropriate APCs: if the A → B chimaeras did contain T helper cells able to recognize SRBC associated with the donor B cell MHC, there could be a problem with the activation of these T cells. Although the chimaeras contain professional APCs of the host MHC type, they have no APCs capable of presenting SRBC associated with donor MHC. They do, of course, contain donor B cells, but if resting B cells cannot deliver the appropriate triggering signals, then resting T cells specific for SRBC + donor MHC will not become activated and no B cell responses would ensue. If this were true, providing the A → B chimaeras with

donor type APCs should restore the response. To test for this, we injected irradiated donor type spleen cells along with SRBC. Table 2 shows that the A $\rightarrow$ B chimaeras become fully responsive when the antigen is given together with a source of irradiated 'A'-type APCs identical to the B-cell donor and remain unresponsive if injected with host-type APCs.

Before concluding that the restored responses were due to the antigen presenting capacity of the injected spleen cells, we asked whether they might instead be the result of nonspecific allogeneic effects. Allogeneic recognition by T cells, even when irradiated, has long been known to result in both positive²² and negative²³ immune responses. Therefore, our 'A'-type injected spleen cells might well have recognized host MHC antigens and consequently induced a spurious positive response. We tested this by injecting APCs from various sources, only some of which could give rise to allogeneic effects (Table 3). We found that A $\times$ B hybrid spleen cells (group ii), which can recognize neither donor nor host, are perfectly capable of restoring the responses of A $\rightarrow$ B chimaeras, showing that the response is restored in the absence of allogeneic recognition. We also injected spleen cells capable of recognizing either donor or host cells, or both (groups iii-v). In every case we saw a response only when the injected APCs shared MHC antigens with the donor B cells, indicating that an allogeneic effect is neither necessary nor sufficient to generate a T-dependent B-cell response in the chimaeras.

Thus, the chimaeras are not defective because they lack T cells able to recognize antigens associated with the donor MHC. Nor are they defective because the reconstituting B cells are not able to display antigen and receive help to make T-dependent responses. The prime defect is the lack of donor APCs. Although the donor B cells can display antigen, they cannot provide the appropriate triggering signals to activate resting T cells and initiate an immune response.

What then of the *in vitro* and *in vivo* evidence showing that B cells can trigger T cells¹⁻⁶? These studies, having been done with T-cell lines, clones and hybridomas, point very clearly to a possible difference between activated T cells and their virgin precursors. Obviously self tolerance should operate at the level of the resting T cell and, although unresponsiveness to most self antigens can be maintained by deletion of the specific T cells or sequestration of the antigen into immunologically privileged sites, neither tactic can be used to generate tolerance of idiotypes. Being part of the immune system itself, the idiotypes cannot be sequestered and, being heterogeneous, they potentially make up such a large set of antigens that tolerance by deletion could cripple the system. It cannot be argued that the concentration of each individual idio type is too low to be either tolerogenic or stimulatory because the concentration of idio type on the surface of each B cell is as high as, or higher than, other surface molecules^{24,25}. Nor can it be argued that endogenous immunoglobulins are not processed and displayed because both anti-allotypic²⁶ and anti-idiotypic²⁷ T-cell lines exist which can respond to a B-cell tumour displaying its own antibody²⁸. The immune system's solution to the problem is simple; it does not allow resting B cells to trigger resting T cells, thus ensuring the maintenance of unresponsiveness to self idiotypes by preventing primary T-cell activation.

The situation may, however, be quite different for a secondary T-cell response, the efficiency of which could be greatly increased by allowing the reactivation of memory or previously activated T cells by any cell capable of displaying the right antigen. B cells are excellent for this purpose, since they are able to pick up and display minute quantities of antigen⁴. Thus both resting and activated B cells should be (and are) able to re-stimulate T cells. But neither resting nor activated B cells should be able to activate a population of truly resting, virgin T cells.

We thank Anna Alanen, Harald von Boehmer, David Gray, Antonio Lanzavecchia, Charles Mackay, Fritz Melchers, Ronald Palacios, Richard Pink, Mike Ratcliffe, Uwe Staerz, Charley

Steinberg and Jean-Claude Weill for critical comments, Viktor Hasler, Brigit Kugelberg, Bente Walch and Brigitte Riwar for assistance with the experiments, and Carolyn Harley for expert typing. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche & Co. Ltd.

Received 30 March; accepted 2 June 1988.

1. Chesnut, R. W. & Grey, H. M. *Adv. Immun.* **39**, 51-92 (1986).
2. Rock, K. L., Benacerraf, B. & Abbas, A. K. *J. exp. Med.* **160**, 1102-1113 (1984).
3. Ashwell, J. D., De Franco, A. L., Paul, W. E. & Schwartz, R. H. *J. exp. Med.* **159**, 881-905 (1984).
4. Lanzavecchia, A. *Nature* **314**, 537-539 (1985).
5. Tony, H. B. & Parker, D. C. *J. exp. Med.* **161**, 223-241 (1985).
6. Ron, Y. & Sprent, J. *J. Immun.* **138**, 2848-2856 (1987).
7. Toivanen, P., Toivanen, A. & Sorvari, T. *Proc. natn. Acad. Sci. U.S.A.* **71**, 957-962 (1974).
8. Vainio, O., Koch, C. & Toivanen, A. *Immunogenetics*, **19**, 131-139 (1984).
9. Pink, J. R. L. & Lassila, O. *Curr. Topics Microbiol. Immun.* **135**, 57-64 (1987).
10. Lehtonen, L., Vainio, O., Eerola, E. & Toivanen, P. *Transplantation* **40**, 398-405 (1985).
11. Pink, J. R. L., Droegge, W., Hala, K., Miggiano, V. C. & Ziegler, A. *Immunogenetics* **5**, 203-216 (1977).
12. Briles, W. E. *et al. Immunogenetics* **15**, 441-450 (1982).
13. Gronowicz, E., Coutinho, A. & Melchers, F. *Eur. J. Immun.* **6**, 588-590 (1976).
14. Matzinger, P. *Nature* **292**, 497-501 (1981).
15. Gorman, J., Hunkapillar, T. & Hood, L. *Cell* **45**, 475-484 (1986).
16. Bevan, M. J. *Nature* **269**, 417-419 (1977).
17. Zinkernagel, R. M. *et al. J. exp. Med.* **147**, 882-896 (1978).
18. Heber-Katz, E. & Wilson, D. B. *J. exp. Med.* **142**, 928-935 (1975).
19. Waldmann, H., Pope, H., Brent, L. & Bighouse, K. *Nature* **274**, 166-168 (1978).
20. Matzinger, P. & Mirkwood, G. *J. exp. Med.* **148**, 84-92 (1978).
21. Wagner, H., Hardt, C., Stockinger, H., Pfizenmeier, K., Bartlett, R. & Röllinghoff, M. *Immun. Rev.* **58**, 95-129 (1981).
22. Katz, D. H. *Transplant. Rev.* **12**, 141-179 (1972).
23. Swain, S. L. & Dutton, R. W. *J. Immun.* **118**, 2262-2268 (1977).
24. Rabellino, E., Colon, S., Grey, H. M. & Unanue, E. R. *J. exp. Med.* **133**, 156-167 (1971).
25. Andersson, J., Lafleur, L. & Melchers, F. *Eur. J. Immun.* **4**, 170-180 (1974).
26. Bikoff, E. & Birshtein, B. K. *J. Immun.* **137**, 28-34 (1986).
27. Bogen, B., Malissen, B. & Haas, W. *Eur. J. Immun.* **16**, 1373-1380 (1986).
28. Bogen, B. & Weiss, S. *Proc. Germinal Center Conference*, (Plenum Press, New York, in the press.)

Regulation of antibody isotype secretion by subsets of antigen-specific helper T cells

Tracy L. Stevens†, Alexis Bossie†, Virginia M. Sanders*, Rafael Fernandez-Botran*, Robert L. Coffman‡, Timothy R. Mosmann‡ & Ellen S. Vitetta*†

* Department of Microbiology and the † Immunology Graduate Program, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235, USA and ‡ DNAX Research Institute, Palo Alto, California 94304, USA

The regulation of the subclass of immunoglobulin secreted by B cells has been studied *in vitro* in polyclonal systems using mitogens, such as lipopolysaccharide (LPS), to bypass the requirement for cognate interaction between antigen-specific T and B cells. In these systems, interleukin-(IL)-4 induces the secretion of IgG₁ (ref. 1) and IgE (ref. 2); IL-5 enhances the secretion of IgA^{3,4}, and interferon- γ (IFN- γ) enhances the secretion of IgG_{2a} (ref. 5). Clones of murine T_H cells can be divided into two subsets, T_{H1} and T_{H2} (ref. 6). Both subsets synthesize IL-3 and granulocyte-macrophage colony-stimulating factor (GM-CSF), but only T_{H1} clones produce IL-2, IFN- γ , and lymphotoxin (LT) and T_{H2} clones produce IL-4 and IL-5 (ref. 7). We have examined the role of clones of antigen-specific T_{H1} and T_{H2} cells in the regulation of the subclasses of IgG antibody secreted by antigen-specific B cells. Our results show that both types of T_H cells induce the secretion of IgM and IgG₃, whereas clones of T_{H1} and T_{H2} cells specifically induce antigen-specific B cells to secrete IgG_{2a} and IgG₁, respectively. We also demonstrate that regulation of commitment to the secretion of a particular IgG isotype occurs in two distinct stages: cognate interaction between T and B cells and interaction between T-cell-derived lymphokines and B cells.

B cells specific for trinitrophenyl (TNP-ABC) were derived from the spleens of normal mice^{8,9}. Of the purified TNP-ABC

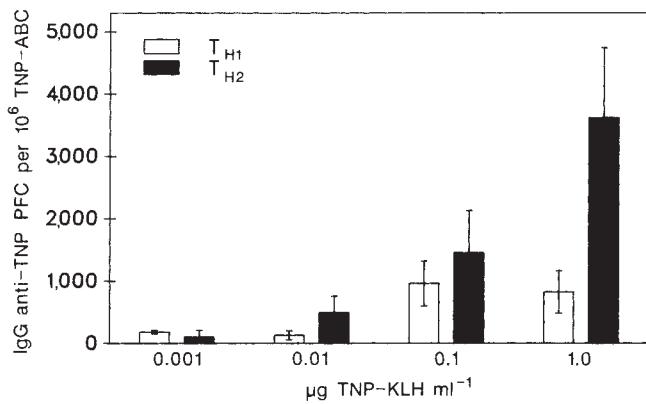


Fig. 1 Antigen-specific PFC responses are generated by B cells activated with specific antigen and T_{H1} (HDK-1) or T_{H2} (T-286) cells. TNP-ABC were prepared as previously described from the spleens of unprimed BALB/c mice^{9,22}. Briefly, B cells were mixed with TNP-haptenated horse red blood cells (TNP-HRBC). The resulting rosettes were separated from non-rosetted cells by centrifugation over two successive discontinuous Percoll gradients. TNP-HRBC and surface immunoglobulin were cleaved off the TNP-ABC by mild trypsin-pronase treatment followed by passage over Ficol to separate the HRBC from the TNP-ABC. The TNP-ABC were incubated in 25 cm² flasks overnight at 37 °C in a humidified atmosphere of 83% N₂, 10% CO₂ and 7% O₂ to allow reexpression of surface immunoglobulin. After 18–24 h of incubation, the TNP-ABC were harvested and plated at 5 × 10⁴ cells per well in 96-well microtitre plates with 2 × 10⁴ T_{H1} (HDK-1)⁷ or T_{H2} (T-286)¹² cells per well and 0.001–1.0 µg ml⁻¹ of TNP-KLH. Both T_H clones were derived from BALB/c mice. The total volume per well was 200 µl. Cultures were incubated at 37 °C as described above for six days. Cells were then washed and harvested and IgG PFC assays were performed using TNP-SRBC in a microscope slide modification of the haemolysis in agar technique. IgG plaques were determined by subtracting direct plaques (IgM only) from indirect plaques (IgM + IgG) which were developed using rabbit anti-mouse-γ antiserum in the assay. At these antigen concentrations, the non-cognate (KLH only) responses ranged from 0% to 7% of the cognate response for T_{H1} and 0% to 15% for T_{H2}. This figure represents the mean ± s.e.m. from three experiments. (□) TNP-ABC + TNP-KLH + T_{H1}; (■) TNP-ABC + TNP-KLH + T_{H2}.

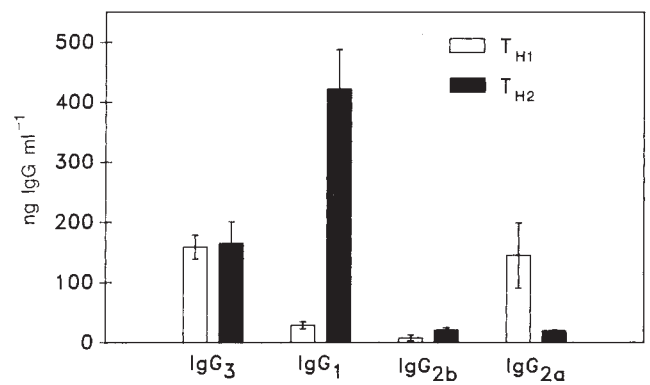


Fig. 2 Isotypes of IgG produced by B cells activated with specific antigen and T_{H1} (HDK-1) or T_{H2} (T-286) cells. Cells were prepared and cultures were set up as described in the legend to Fig. 1 using 1 µg ml⁻¹ of antigen. At the end of the incubation period, supernatants from triplicate wells were pooled and concentrations of the four IgG isotypes were determined using a solid-phase radioimmunoassay (RIA). The RIA was carried out as follows; affinity-purified goat anti-mouse γ-chain antibody was bound to 96-well flexible polyvinyl chloride plates at 5 µg ml⁻¹ in phosphate-buffered saline containing 10 mM sodium azide (PBS-NaN₃) for 4 h at room temperature. The plates were washed and blocked with PBS-NaN₃ containing 1% fetal bovine serum (PBS-FBS) and stored at 4 °C. Supernatants to be tested were diluted in PBS-FBS and incubated on the goat anti-γ-coated plates for 4 h at room temperature, washed, and developed with ¹²⁵I-labelled affinity-purified goat anti-isotype antibody for 4 h at room temperature. Wells were washed, cut and counted. Antibody concentrations were determined from standard curves using myeloma proteins of the four IgG isotypes. This figure represents the mean ± standard error (s.e.) of triplicate determinations from three experiments. (□) TNP-ABC + TNP-KLH + T_{H1}; (■) TNP-ABC + TNP-KLH + T_{H2}. The levels of IgM in culture SN for both T_{H1} and T_{H2} and TNP-ABC + TNP-KLH exceeded 1 µg ml⁻¹. No IgM or IgG was secreted in the absence of T_H cells.

>75% bind TNP and >90% of the cells are dense, resting (G₀) cells. Previous studies using TNP-ABC at limiting dilution have shown that 1/267 secrete IgG (300 ng per 10⁵ cells) and 1/15 secrete IgM antibody (2–3 µg per 10⁵ cells) in response to the thymus-dependent (TD) antigen, TNP-keyhole limpet haemocyanin (KLH) and KLH-specific T cells^{10,11}. The two T_H cell types used to activate the TNP-ABC are both KLH-specific and I-A^d-restricted, but one (HDK-1) is a T_{H1} (ref. 7) and the other (T-286) is a T_{H2} (ref. 12). As expected, when the T_{H1} cell types were cultured with TNP-KLH and antigen-presenting cells (APC), IL-2 and IFN-γ, but not IL-4 were secreted⁷ and when the T_{H2} cells were stimulated in the same manner, IL-4, but not IL-2 or IFN-γ, was secreted¹² (Table 1). As confirmed by dot-blot hybridization analysis of RNA, cells from the T_{H1} clone do not synthesize messenger RNA for IL-4 or IL-5 and cells from the T_{H2} clone do not synthesize mRNA for IL-2 or IFN-γ^{7,12}.

As shown in Fig. 1, when the TNP-ABC were cultured for six days with T_{H1} or T_{H2} cells in the presence of TNP-KLH, both types of T_H cells induced IgG anti-TNP plaque-forming cells (PFC). Furthermore, as shown in Fig. 2, the T_{H1} cells induced the secretion of IgG₃ and IgG_{2a}, but little or no IgG₁. In contrast, the T_{H2} cells induced the secretion of IgG₃ and IgG₁, but not IgG_{2a}. Neither of the two T_H clones induced the secretion of IgG_{2b} and both induced the secretion of >1 µg ml⁻¹ IgM.

The results presented in Figs 1 and 2 indicate that T_{H1} (HDK-1) and T_{H2} (T-286) cells can both provide help to TNP-ABC.

Thus, the T_{H1} clone, HDK-1, unlike many other T_{H1} clones^{13,14}, can help B cells, and the pattern of IgG isotypes secreted differs from that seen with the T_{H2} clone, T-286.

Earlier studies by Layton *et al.*¹⁰ demonstrated that IL-4 induced the production of IgG₁ in B cells cultured with LPS and that this was due to an increase in the precursor frequency of IgG₁-secreting cells. Furthermore, precursors lacked surface IgG (sIgG), suggesting that IL-4 was a switch factor. The induction of IgG secretion in the experiments described here may occur *via* the same mechanism, that is, switching from IgM/IgG₃ to IgG₁ or IgG_{2a} at the clonal level rather than expansion of pre-existing sIgG₁⁺ or sIgG_{2a}⁺ cells. We have not, however, formally excluded the possibility either that the IgG responses reported here result from the activation of sIgG⁺ rather than sIgG⁻ cells (even though the mice were not primed), or that they require an increase in precursor frequency. The answers to these questions will involve experiments using limiting-dilution analysis.

Because it has been reported that, in polyclonal systems, IFN-γ (produced by T_{H1} cells) and IL-4 (produced by T_{H2} cells) induce B cells to secrete IgG_{2a} and IgG₁ respectively⁵, we tested the effects of adding monoclonal anti-IFN-γ⁷ to cultures containing T_{H1} cells and monoclonal anti-IL-4 (ref. 15) to cultures containing T_{H2} cells. In 9 of 18 experiments, antibody secretion was blocked by 20–80% (mean (±s.d.) = 52% (±21%)) blocking by anti-IL-4 and 41% (±21%) blocking by anti-IFN-γ. In the remaining experiments, neither antibody could block IgG_{2a} and IgG₁ responses (data not shown). That these antibodies do not consistently or completely block the secretion of IgG₁ and IgG_{2a} could not be attributed to insufficient amounts of the antibodies

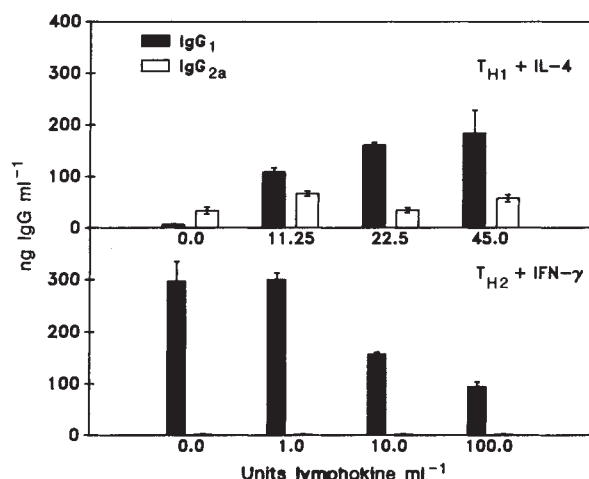


Fig. 3 The influence of exogenous lymphokines on the secretion of IgG₁ and IgG_{2a} by TNP-ABC cultured with T_{H1} (top) or T_{H2} (bottom) cells. In the top panel, TNP-ABC were incubated with T_{H1} cells and TNP-KLH ± increasing concentrations of IL-4. Supernatants were collected at day 6 and assayed for IgG₁ (■) and IgG_{2a} (□). In the lower panel, TNP-ABC were incubated with T_{H2} cells and increasing amounts of IFN-γ. SNs were collected on day 6 and assayed for IgG₁ (■) and IgG_{2a} (□). Both panels represent the mean ± standard error of three separate experiments. The IgM responses ranged from 500 to 1500 ng ml⁻¹ in the absence of lymphokines and were increased 3–10-fold by the addition of IL-4 and decreased up to 80% by the addition of IFN-γ.

Table 1 Lymphokines produced by two types of T_H cell lines after antigenic stimulation

Cell lines	Activity (U ml ⁻¹)		
	IL-2	IL-4	IFN-γ
HDK-1 (T _{H1})	25	<0.1	23.5
T-286 (T _{H2})	<0.1	37.5	<0.01

Culture supernatant from antigen-stimulated T_{H1} (HDK-1)⁷ and T_{H2} (T-286)¹² cells were collected after 24 h of incubation of the T cells at a density of 5×10^5 cells ml⁻¹ in the presence of KLH (50 μg ml⁻¹) and irradiated (3,300 rad) splenic adherent cells (SAC) from BALB/c mice. After filtration (0.22 μm), the T-cell supernatants were tested for the presence of IL-2, IL-4 and IFN-γ as previously described^{12,21}. Briefly, IL-2- and IL-4 activities were measured by their ability to support the proliferation of HT-2 cells¹² in the absence and presence of the anti-IL-2 (S4B6)⁶ and anti-IL-4 (11B11)¹⁵ monoclonal antibodies which block IL-2 and IL-4-mediated proliferation, respectively. IFN-γ activity was measured by the ability of the IFN-γ contained in the supernatants to inhibit the proliferation of the B lymphoma cell line, WEHI-279 (ref. 21), coupled to its reversibility by the anti-murine IFN-γ antibody (XMG1.2)⁷.

Activities are expressed as units (U) ml⁻¹. One unit is defined as the amount of any particular lymphokine causing 50% of the maximal response of a recombinant lymphokine standard in the previously described assays. Supernatants from control cultures (T cells only, T cells + APC, T cells + KLH, or APC + KLH) did not show any detectable levels of these lymphokines.

in the cultures as the addition of a total of 15 μg ml⁻¹ of either antibody on days 0, 1 and 4 did not cause inhibition (data not shown). The failure to block IgG secretion could be due to close physical interaction between T and B cells^{16,17} such that IFN-γ and IL-4 are transferred between the cells in a manner that is refractory to antibody blockade or to the presence of other IgG₁- or IgG_{2a}-inducing factors produced by cells in the cultures.

As IL-4 (produced by the T_{H2} clone) induces mitogen-stimulated B cells to secrete IgG₁, but not IgG_{2a}, we next determined whether IL-4 added to cultures containing T_{H1} cells could induce an increase in the secretion of IgG₁. As shown in Fig. 3 (top panel), addition of IL-4 induced increased secretion of IgG₁. Although it is not known whether exogenously added IL-4 and endogenously-secreted IFN-γ act on the same or different TNP-ABC, we suggest that the TNP-ABC are different because it is unlikely that cells which have switched to IgG_{2a} would be induced to secrete IgG₁, which is the product of an upstream (and probably deleted) C_H gene; and the IgG_{2a} response does not decrease as the IgG₁ response increases. In any case, in the absence of T_{H1} cells, neither isotype was secreted, indicating that IL-4 alone was insufficient to induce IgG₁ secretion. Surprisingly, the IgG₁ response was induced in the presence of endogenously secreted IFN-γ which is normally suppressive for the IgG₁-inducing effects of IL-4 in polyclonal systems⁵. Therefore, a signal delivered by the T_{H1} cell can prime one or more subsets of TNP-ABC to secrete IgG₁ and/or IgG_{2a} provided that the appropriate lymphokines are present.

In contrast, when IFN-γ was added to cultures containing TNP-ABC, TNP-KLH and cells from the T_{H2} clone, an IgG_{2a} response could not be elicited (Fig. 3, lower panel). Furthermore, the addition of anti-IL-4 (11B11)¹⁵ to cultures had no effect (data not shown). In the same experiments, IFN-γ inhibited the IgG₁-inducing effects of the T_{H2} cells. Several mechanisms may be responsible for the failure of the TNP-ABC to secrete IgG_{2a} when cultured with cells from the T_{H2} clone and IFN-γ. First, the T_{H2} cell itself, or a lymphokine which it secretes, may either prevent or down-regulate the activation of precursors of IgG_{2a}-secreting cells; second, lymphokines derived from the T_{H2}

clone may inhibit the activities of the IFN-γ or other IgG_{2a}-inducing factors; or third, cells from the T_{H2} clone may be unable to activate IgG_{2a} precursors because of the absence of an appropriate signal generated during T_{H2} cell/B cell interaction. In an attempt to distinguish between these possibilities, IFN-γ was added to the TNP-ABC 24–48 h before T_{H2} cells and antigen were added. Although a modest IgG_{2a} response was seen in a few experiments, the vast majority of the experiments were negative. Thus, even though IFN-γ can 'program' B cells to secrete IgG_{2a} in polyclonal systems⁵, this does not occur using TNP-ABC and T_{H2} cells. Therefore, we cannot exclude any of the possibilities described above.

In conclusion, we have demonstrated that cells of two T_H types, which are specific for epitopes on the same antigen and are restricted by the same I-region haplotype, induce a population of antigen-specific B cells to secrete either IgG₁ (T_{H2} cells) or IgG_{2a} (T_{H1} cells). Our findings further demonstrate that the regulation of isotype secretion by cells of a T_{H1} clone occurs at two levels: one involves T_H cell/B cell interaction, the other involves T-cell-derived lymphokines. These results indicate that the regulation of isotype switching *in vivo* will depend upon the types of T_H cells that are activated and whether lymphokines are transferred between T and B cells in a manner that is refractory to anti-lymphokine antibodies. Such considerations are pertinent in designing strategies to use lymphokines or anti-lymphokine antibodies *in vivo* to induce the secretion of immunoglobulin isotypes which will subserve different biological functions and hence induce an optimal immune response to a particular type of pathogen, for example viruses¹⁸, bacteria¹⁹ or parasites²⁰.

We thank Ms L. Trahan, Ms S. Shanahan, Ms M.-M. Liu, Ms T. R. Wilson, Ms P. May and H. Cherwinski for expert technical assistance and Ms G. A. Cheek for secretarial assistance.

Received 14 March; accepted 24 May 1988.

- Isakson, P., Pure, E., Vitetta, E. S. & Krammer, P. H. *J. exp. Med.* **155**, 734–748 (1982).
- Coffman, R., Ohara, J., Bond, M., Carty, J., Zlotnik, A. & Paul, W. E. *J. Immun.* **136**, 4538–4541 (1986).
- Coffman, R. L., Shrader, B., Carty, J., Mosmann, T. R. & Bond, M. W. *J. Immun.* **139**, 3685–3689 (1987).

4. Bond, M. W., Shrader, B., Mosmann, T. & Coffman, R. L. *J. Immun.* **139**, 3691-3696 (1987).
5. Snapper, C. M. & Paul, E. S. *Science* **236**, 944-947 (1987).
6. Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A. & Coffman, R. L. *J. Immun.* **136**, 2348-2357 (1986).
7. Cherwinski, H. M., Schumacher, J. H., Brown, K. D. & Mosmann, T. R. *J. exp. Med.* **166**, 1229-1233 (1987).
8. Vitetta, E. S. *et al. Immun. Rev.* **99**, 193-239 (1987).
9. Myers, C. D., Sanders, V. M. & Vitetta, E. S. *J. immun. Meth.* **92**, 45-57 (1986).
10. Layton, J. E., Uhr, J. W., Vitetta, E. S. & Krammer, P. H. *J. exp. Med.* **160**, 1850-1863 (1984).
11. Yefenof, E. *et al. J. Immun.* **135**, 3777-3784 (1985).
12. Fernandez-Botran, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9689-9693 (1986).
13. Killar, L., MacDonald, G., West, J., Woods, A. & Bottomly, K. *J. Immun.* **138**, 1674-1679 (1987).
14. Boom, W. H., Liano, D. & Abbas, A. *J. exp. Med.* **167**, 1350-1363 (1988).
15. Ohara, J. & Paul, W. E. *Nature* **315**, 333-336 (1985).
16. Sanders, V. M., Snyder, J. M., Uhr, J. W. & Vitetta, E. S. *J. Immun.* **137**, 2395-2404 (1986).
17. Kupfer, A., Swain, S. L., Janeway, C. A. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6080-6083 (1986).
18. Coutelier, J.-P., van der Logt, J. T. M., Heesen, F. W. A., Warnier, G. & Van Snick, J. *J. exp. Med.* **165**, 64-69 (1987).
19. Perlmuter, R. M., Hansburg, D., Butler, D. E., Nicolotti, R. A. & Davie, J. M. *J. Immun.* **121**, 512-566 (1978).
20. Jarrett, E. E. & Bazin, H. *Nature* **251**, 613-614 (1974).
21. Reynolds, D. S., Boom, W. H. & Abbas, A. K. *J. Immun.* **139**, 767-773 (1987).
22. Snow, E. C., Vitetta, E. S. & Uhr, J. W. *J. Immun.* **130**, 607-613 (1983).

Cytotoxic T cells specific for the circumsporozoite protein of *Plasmodium falciparum*

Sanjai Kumar*, Louis H. Miller*, Isabella A. Quakyi*, David B. Keister*, Richard A. Houghten†, W. Lee Maloy‡, Bernard Moss§, Jay A. Berzofsky|| & Michael F. Good*

* Laboratory of Parasitic Diseases, ‡ Biological Resources Branch, and § Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases; || Metabolism Branch, National Cancer Institute; National Institutes of Health, Bethesda, Maryland 20892, USA; † Department of Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, California 92037, USA

Malaria is initiated by the inoculation of a susceptible host with sporozoites from an infected mosquito. The sporozoites enter hepatocytes and develop for a period as exoerythrocyte or hepatic stage parasites¹. Vaccination with irradiated sporozoites can provide protective immunity¹ and a recent study² shows that this can also be conferred by immunization with a recombinant salmonella expressing only the circumsporozoite protein that normally covers the sporozoites. Protection against infection is likely to be mediated by cytotoxic CD8⁺ cells, as depletion of CD8⁺ T cells in a sporozoite-immunized animal can completely abrogate immunity^{3,4}. Here we demonstrate directly the existence of CD8⁺ cytotoxic T lymphocytes (CTL) that recognize the circumsporozoite protein. B10.BR mice immunized with sporozoites or with recombinant vaccinia virus expressing the CS protein of *Plasmodium falciparum* contain CTL that specifically kill L cell fibroblasts transfected with the gene encoding the same CS protein. The peptide epitope from the CS protein that is recognized by CTL from this strain of mice is from a variant region of the protein.

To study CTL against the CS protein, mouse L cells (H-2^k) were transfected with the gene encoding the CS protein of the 7G8 clone of *P. falciparum*⁵. Two clones were derived by limiting dilution, and expression was confirmed by immunofluorescence which revealed cytoplasmic but not surface expression. To generate putative CTL, B10.BR mice (H-2^k) were inoculated intravenously with 10⁷ PFU of a recombinant vaccinia virus containing the same CS gene^{6,7}. Spleen cells from immunized mice were then stimulated *in vitro* for seven days with a transfected L cell clone, based on the method of Takahashi *et al.*⁸. Cultured cells were then assessed for lytic activity against either CS-transfected L cells or other control target cells. We observed specific killing of the CS gene-transfected L cells indicating that

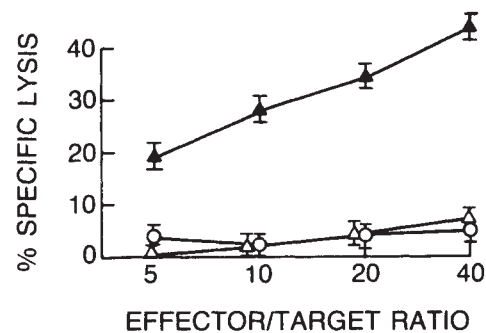


Fig. 1 The CS protein is a target for CTL. B10.BR mice (H-2^k) were immunized intravenously with 10⁷ PFU of a CS-recombinant vaccinia virus^{6,7}. Three weeks later, spleens were taken and 5 × 10⁶ spleen cells incubated with 2 × 10⁵ CS-transfected L cells treated with mitomycin C (50 µg ml⁻¹, 30 min) in 2 ml RPMI 1640 medium containing 10% fetal calf serum, 5 × 10⁻⁵ M 2-mercaptoethanol and penicillin and streptomycin. After 7 days' culture, cells were harvested, counted and incubated for 6 h at various effector/target ratios with 5,000 ⁵¹Cr-labelled CS-transfected L cells, ▲; pSV2-neo transfected L cells (L-28), △; or untransfected, uninfected L cells, ○. Per cent specific lysis was determined as:

$$\frac{(\text{experimental c.p.m.} - \text{medium control c.p.m.})}{(\text{detergent-released c.p.m.} - \text{medium control c.p.m.})} \times 100\%$$

The 'spontaneous release' values (medium control c.p.m./detergent-released c.p.m.) for all targets for all the experiments reported in this paper were very similar with a mean value of 16.8% ± 1.1% (s.e.m.). Error bars on this and other figures represent ± s.e.m. (triplicate assays). Unstimulated spleen cells from unimmunized mice did not kill target cells.

Methods. Mouse L cells (H-2^k) were transfected with the CS gene of the 7G8 clone of *P. falciparum*⁵ by calcium phosphate precipitation¹⁷. Briefly, the CS gene containing the entire coding sequence was cloned into the EcoRI site of the expression vector pcEXV-3 (ref. 18). This expression vector contains the SV40 enhancer, early promoter and splice signals 5' of the gene insert as well as a polyadenylation signal downstream. The rest of the sequences are derived from pBR322. L cells (5 × 10⁵ cells per 60-mm petri dish) were cotransfected in 1 ml with 10 µg of purified plasmid DNA and 3 µg pSV2neo DNA. The calcium phosphate/DNA precipitate was left on the cells overnight. The following day, medium was replaced with fresh medium and left for an additional 24 h before selection by incubation in medium containing 2 mg ml⁻¹ G418 (Geneticin, Gibco) for 1 week. Finally, stable transfected cells were grown in the presence of 200 µg ml⁻¹ G418. Single colonies of transfectants were obtained by limiting dilution and screened by immunofluorescence for the expression of CS protein.

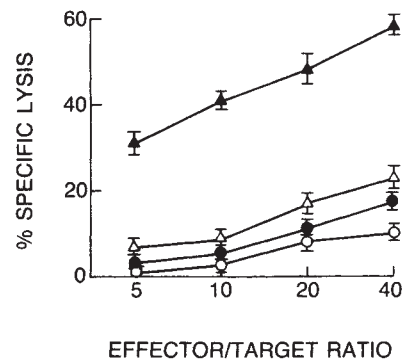


Fig. 2 Sporozoites stimulate CS-specific CTL. B10.BR mice were immunized intravenously with 10⁵ (irradiated, 10,000 rad) *P. falciparum* sporozoites (7G8 clone) prepared as previously described¹⁹ (▲, △). Control mice were not immunized (●, ○). After 3 weeks, mice were sacrificed and spleens taken. Spleen cells were then stimulated *in vitro* with CS-transfected L cells as described in Fig. 1. After 6 days, killing of CS-transfected L cells (▲, ●) and untransfected L cells (△, ○) was measured.

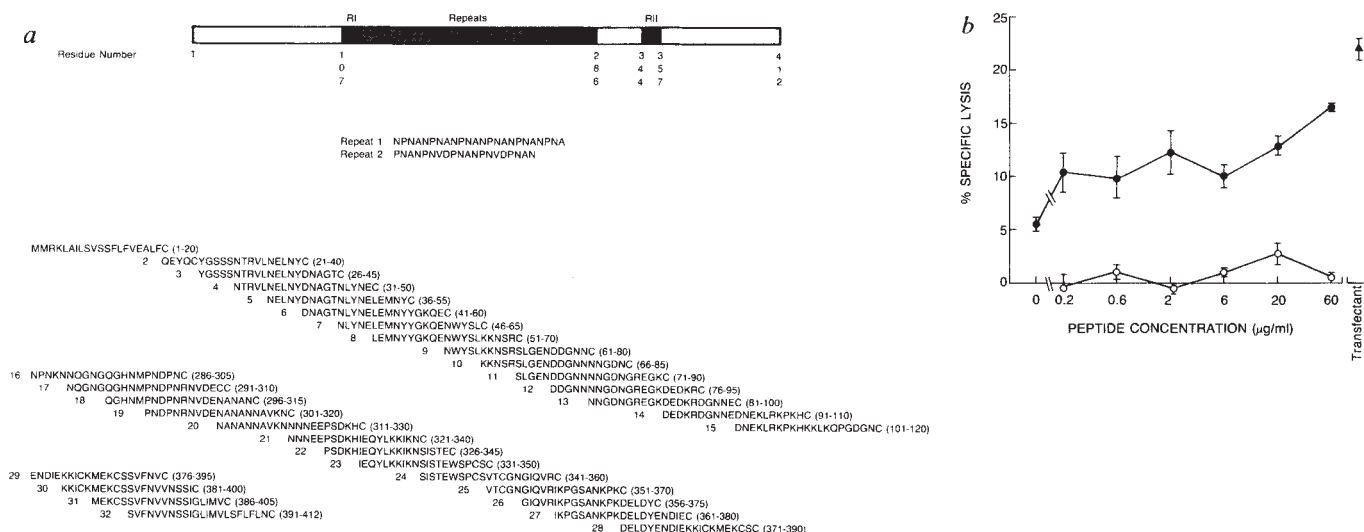


Fig. 3 *a*, Schematic illustration of the CS protein and description of peptides. RI (Region I) and RII (Region II) refer to areas of strong homology between CS proteins from different species of malaria⁵. The sequences of the 34 peptides tested were taken from the 7G8 clone⁵. Each peptide has a carboxyl terminal cysteine which is not part of the protein sequence. The numbers in brackets refer to the residue number. Peptides were produced as described²⁰ and desalted prior to use. They were checked for purity by reversed-phase HPLC. The peptides were not toxic to target cells. *b*, Definition of the CTL epitope on the CS protein. Spleen cells from CS-vaccinia-immunized B10.BR mice were stimulated *in vitro* with CS-transfected L cells as described in Fig. 1. Putative CTL were then incubated with untransfected L cells at an effector/target ratio of 10:1 for 6 h in the presence of various concentrations of peptide 28 (●). There was no specific lysis of L cells by peptide 28 in the absence of CTL (○), showing that the peptide was not toxic to these cells. The killing of the transfected L cell was also measured, at an effector/target ratio of 10:1 (▲). Anti-CS-specific CTL did not kill L cells in the presence of any of the other 33 peptides at concentrations of 1, 10, 20 or 100 $\mu\text{g ml}^{-1}$. Peptide 28 did not sensitize L cells to killing by irrelevant allospecific CTL (C57B1/10 anti-BALB/c, data not shown).

the CS protein did contain one or more epitopes recognized by CTL (Fig. 1 and replicate data not shown). These CTL also recognized the other CS-transfected L cell clone as well as CS-recombinant vaccinia-infected L cells, but not wild-type vaccinia-infected L cells (data not shown). CS-specific CTL were sensitive to anti-CD8 monoclonal antibody and complement. Following incubation of CS-specific CTL with anti-CD8 (19/178 clone⁹) followed by complement, there was 3.8% specific lysis of CS-transfected targets at an effector/target ratio of 10:1, compared with a specific lysis of 29% when CTL were incubated with complement alone. Lysis of untransfected L cells was 3.2% following complement treatment, and 0.12% following treatment with anti-CD8 and complement. These data show that the CS-specific cytotoxicity was mediated by CD8⁺ cells and, in keeping with previous findings^{10,11}, that it did not require cell-surface expression of the CS protein.

Whereas stimulation of CTL by virus-encoded proteins is well documented, there has been no direct demonstration of CTL by malaria parasites to date. We were therefore interested to know whether CS-specific CTL can be induced by sporozoites. B10.BR mice were immunized intravenously with freshly prepared 7G8 sporozoites, and three weeks later spleen cells were taken and stimulated *in vitro* with CS gene-transfected L cells. After seven days' culture, specific lytic activity could be demonstrated against transfected L cells (Fig. 2). Spleen cells from unimmunized B10.BR mice could not be stimulated to generate CS protein-specific lytic activity (Fig. 2). Thus sporozoites can stimulate CS-specific CTL *in vivo*.

To determine the epitope recognized by CS-specific CTL, a series of overlapping peptides spanning the entire CS protein were constructed (Fig. 3*a*, and ref. 12). CTL were then incubated with L cells in the presence of each of the peptides, and lysis determined by ⁵¹Cr release¹³. Peptides were tested at concentrations of 1, 10, 20 and 100 $\mu\text{g ml}^{-1}$. Only one peptide (#28, residues 371–390) was recognized by CTL and for this peptide a dose-response study demonstrated significant killing at concentrations as low as 0.2 $\mu\text{g ml}^{-1}$ (Fig. 3*b*). Specific killing by

B10.BR anti-CS protein CTL in the presence of this peptide was not, however, as great as killing of the transfected L cells, possibly indicating that peptide 28 does not contain the full CTL recognition site. Thus, peptides 370–390 and 368–390 were made (lacking the additional carboxy-terminal cysteine present on peptide 28) and tested for recognition by B10.BR anti-CS protein CTL. L cells incubated with peptide 368–390 were killed

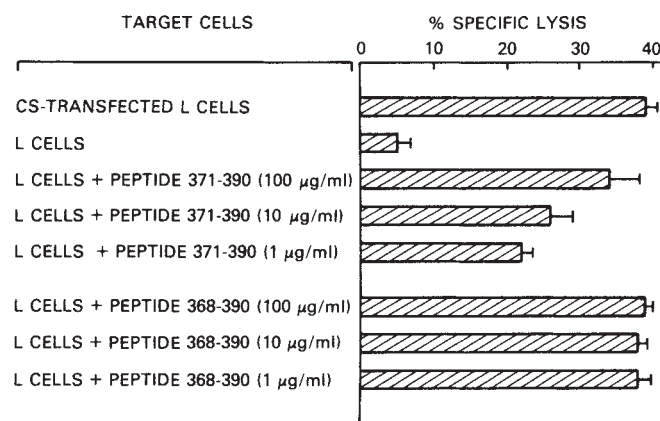


Fig. 4 The CTL epitope is located within residues 368–390. Spleen cells from CS-vaccinia immunized B10.BR mice were stimulated *in vitro* with CS-transfected L cells as described in Fig. 1. Cultured cells were then incubated at an effector/target ratio of 30:1 for 5 h in the presence of various targets: CS-transfected L cells; untransfected L cells; untransfected L cells incubated overnight with peptide (as indicated) at 100 $\mu\text{g ml}^{-1}$, then washed and incubated with peptide at 100, 10, or 1 $\mu\text{g ml}^{-1}$ during the cytotoxicity assay. We found that overnight incubation with peptide increased the degree of specific killing (not shown). Similar results were obtained at effector/target ratios of 10:1. Peptides 370–390 and 368–390 were not toxic to L cell and peptide 368–390 was not recognized by vaccinia-specific CTL (not shown).

at the same level as CS-transfected L cells and more efficiently than L cells incubated with peptide 370–390 (Fig. 4), indicating that the full T cell epitope was contained within residues 368–390. The observation that CTL recognized peptide 368–390 to the same degree as the transfectant gives a further indication that B10.BR CS-specific-CTL recognize only this single epitope on the CS protein. Peptide 368–390, however, represents a polymorphic region of the molecule^{5,14–16}. This peptide epitope was also recognized by CTL generated from sporozoite-immunized mice but not by spleen cells from non-immunized mice stimulated by the transfectant in either the presence or absence of interleukin-2 (data not shown).

Our data demonstrate that malaria sporozoites can induce CS-protein-specific CTL, and that in B10.BR mice the CTL response is restricted to a single peptide epitope from a variant region of the protein. The implications of these results for the development of a malaria vaccine for use in humans are not encouraging. Although T cell responsiveness shows some variation, both with species and MHC haplotype, and human CTL epitopes on the CS protein have not yet been identified, studies of CD4⁺ T cell epitopes on the CS protein suggest that those of mouse and human closely overlap and that they are mostly found in variant regions of the molecule (refs 7, 12 and M.F. Good *et al.*, unpublished observations). It seems likely that selective pressure for evasion of T cell recognition has led to the restricted number and polymorphic nature of T cell recognition sites on the CS protein. If human CTL recognition of the CS protein does prove to be restricted to variant epitopes, it

may be critical for vaccine development to locate other sporozoite or hepatic stage proteins which contain non-polymorphic CD8⁺ T cell recognition sites.

We thank F. A. Neva, R. N. Germain, H. Takahashi, N. Cooper, R. Wirtz, W. E. Biddison and W. Weiss for encouragement and help and Wilma Davis for editorial assistance. I.A.Q. received support from United Nations Development Program/World Bank/WHO Special Program for Research and Training in Tropical Diseases.

Received 4 March; accepted 7 June 1988.

1. Nussenweig, R. & Nussenweig, V. *Phil Trans. R. Soc. B* **307**, 117–128 (1984).
2. Sadoff, J. C. *et al. Science* **240**, 336–338 (1988).
3. Schofield, L. *et al. Nature* **330**, 664–666 (1987).
4. Weiss, W. R., Sedegah, M., Beaudoin, R. L., Miller, L. H. & Good, M. F. *Proc. natn. Acad. Sci. U.S.A.* **85**, 573–576 (1988).
5. Dame, J. B. *et al. Science* **225**, 593–599 (1984).
6. Cheng, K.-C. *et al. in Vaccines 86* (ed. Brown, F., Chanock, R. M. & Lerner, R. A.) 165–168 (Cold Spring Harbor, New York, 1986).
7. Good, M. F. *et al. Science* **235**, 1059–1062 (1987).
8. Takahashi, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 3105–3109 (1988).
9. Hammerling, G. J., Hammerling, U. & Flaherty, L. *J. Exp. Med.* **150**, 108–116 (1979).
10. Townsend, A. R. M., McMichael, A. J., Carter, N. P., Huddleston, J. A. & Brownlee, G. G. *Cell* **39**, 13–25 (1984).
11. Townsend, A. R. M., Bastin, J., Gould, K. & Brownlee, G. G. *Nature* **324**, 575–577 (1986).
12. Good, M. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 1199–1203 (1988).
13. Townsend, A. R. M. *et al. Cell* **44**, 959–968 (1986).
14. Lockyer, M. J. & Schwarz, R. T. *Molec. Biochem. Parasitol.* **22**, 101–108 (1987).
15. del Portillo, H. A., Nussenweig, R. S. & Enea, V. *Molec. Biochem. Parasitol.* **24**, 289–294 (1987).
16. de la Cruz, V. F. *et al. J. Immun.* (in the press).
17. Wigler, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1373–1376 (1979).
18. Miller, J. & Germain, R. N. *J. exp. Med.* **164**, 1478–1489 (1986).
19. Walliker, D. *et al. Science* **236**, 1661–1666 (1987).
20. Houghten, R. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5131–5135 (1985).

Deactivation of macrophages by transforming growth factor- β

Shohko Tsunawaki*‡, Michael Sporn†, Aihao Ding* & Carl Nathan*

* Beatrice and Samuel A. Seaver Laboratory, Division of Hematology–Oncology, Department of Medicine, Cornell University Medical College, New York 10021, USA
† Laboratory of Chemoprevention, National Cancer Institute, Bethesda, Maryland 20892, USA

Macrophage activation—enhanced capacity to kill, in a cell that otherwise mostly scavenges—is essential for host survival from infection and contributes to containment of tumours. Both microbes and tumour cells, therefore, may be under pressure to inhibit or reverse the activation of macrophages. This reasoning led to the demonstration of macrophage deactivating factors from both microbes^{1,2} and tumour cells^{3–5}. In some circumstances the host itself probably requires the ability to deactivate macrophages. Macrophages are essential to the healing of wounds and repair of tissues damaged by inflammation. Yet the cytotoxic products of the activated macrophages can damage endothelium, fibroblasts, smooth muscle and parenchymal cells (reviewed in ref. 6). Thus, after an inflammatory site has been sterilized, the impact of macrophage activation on the host might shift from benefit to detriment. These concepts led us to search for macrophage deactivating effects among polypeptide growth factors that regulate angiogenesis, fibrogenesis and other aspects of tissue repair. Among 11 such factors, two proteins that are 71% similar proved to be potent macrophage deactivators: these are transforming growth factor- β 1 (TGF- β 1) and TGF- β 2.

Growth factors were tested over a broad concentration range for their ability to suppress the capacity of activated mouse peritoneal macrophages to release H₂O₂. The capacity to release H₂O₂ is a close biochemical correlate of macrophage activation,

due to the prominent involvement of reactive oxygen intermediates in their antimicrobial function⁷. Activated macrophages were collected four days after intraperitoneal injection of sodium caseinate and plated at $1.2\text{--}1.3 \times 10^5$ per well in 96-well trays in Eagle's minimum essential medium (α -variant) with 10% horse serum (complete medium). Nonadherent cells were washed off at 2 hours and test media added to triplicate wells for the times indicated before the media were flicked out and the plates washed in saline. H₂O₂ release was then measured in the absence of cytokines over 90 min in response to 167-nM phorbol myristate acetate (PMA) by the horseradish peroxidase-catalysed oxidation of fluorescent scopoletin, and related to the protein content in the same wells, as described⁸.

Incubation of macrophages for two days in $0.04\text{--}100\text{ ng ml}^{-1}$ of TGF- β 1 suppressed H₂O₂ release by $86 \pm 11\%$ (mean $\pm$ s.e.m., $n = 83$ in 25 experiments). Macrophages incubated in medium alone released $243 \pm 65\text{ nmol H}_2\text{O}_2$ per mg cell protein per 90 min ($n = 33$ in 25 experiments). Similar results were obtained with TGF- β 2, although ~ 8 -fold higher concentrations were required. The concentrations causing 50% inhibition of H₂O₂ releasing capacity (EC₅₀) were 0.6 pM for TGF- β 1 and 4.8 pM for TGF- β 2 (Fig. 1), suggesting a physiological role for this action of the cytokines. In contrast, the following growth factors suppressed macrophage H₂O₂ releasing capacity by $\sim 21\%$ to 30% , when tested over a $\geq 10,000$ -fold dose range up to 100 ng ml^{-1} (Fig. 1): natural mouse nerve growth factor, natural murine interleukin (IL)-3, recombinant human IL-1 β , natural bovine fibroblast growth factor a and b, natural porcine platelet-derived growth factor, recombinant mouse colony stimulating factor (CSF) for granulocytes and macrophages, recombinant human CSF for granulocytes, and natural mouse epidermal growth factor (which shares receptors with TGF- α).

Deactivation induced by TGF- β 1 or TGF- β 2 was not evident over the first nine hours of incubation, but became half-maximal by 21–23 h at concentrations of $1\text{--}10\text{ ng ml}^{-1}$, and by $\sim 48\text{ h}$ for 0.1 ng ml^{-1} (Fig. 2). This suggests that TGF- β did not deactivate macrophages by triggering their respiratory burst, because exhaustion of respiratory burst capacity by triggering agents is evident immediately after the burst ceases ($\sim 1.5\text{--}3.5\text{ h}$). Moreover, TGF- β 1 at $1\text{--}100\text{ ng ml}^{-1}$ did not elicit H₂O₂ release

‡ Present address: Second Department of Surgery, Teikyo University School of Medicine, 2-11-1 Kaga, Itabashi-ku, Tokyo 173, Japan.

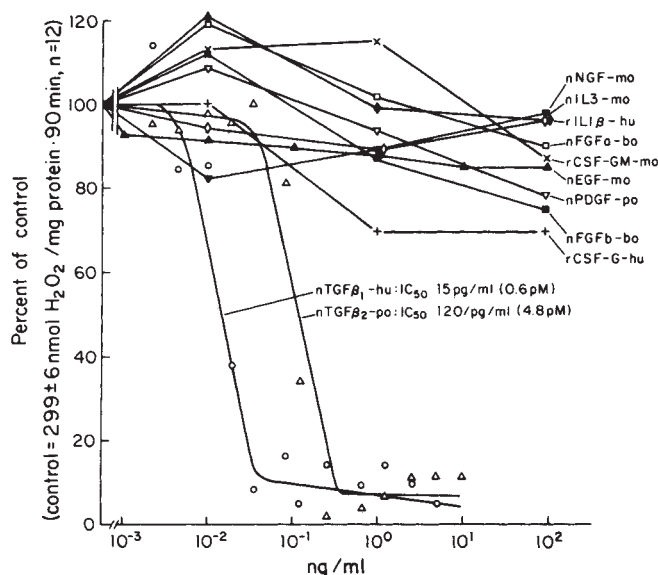


Fig. 1 Suppression of the H_2O_2 releasing capacity of macrophages following incubation for two days in TGF- β_1 and TGF- β_2 , but not any of nine other polypeptide growth factors. Caseinate-elicited mouse peritoneal macrophages were incubated in the indicated concentrations of cytokines before being washed and challenged with 167-nM PMA to trigger secretion of H_2O_2 . One of three similar experiments is shown (two for EGF). For clarity, only means are presented; s.e.m.s for the triplicates averaged 7.5% of the means; n = natural; r = recombinant; suffixes designate species of origin. All reagents were pure, or in the case of IL-3, partially purified. TGF- β_1 was prepared from human platelets as described in ref. 24, except that the urea was removed by desalting on C18 HPLC (high pressure liquid chromatography). TGF- β_2 (ref. 25) was purchased from R&D Systems, Minneapolis. Nerve growth factor (NGF) from mouse submaxillary gland was isolated by Dr Joshua Adler, and provided by Dr Moses Chao, both of Cornell University Medical College. Interleukin-3 (IL-3) (Genzyme, Boston, Massachusetts) was from the conditioned medium of LBRM-33-5A4 T lymphoma cells. Fibroblast growth factors (FGF)-acidic (a) and basic (b) from bovine brain and platelet derived growth factor (PDGF) were from Peninsula Laboratories, Belmont, California. Biologically active epidermal growth factor (EGF) from mouse salivary gland was purified in the laboratory of one of us (M.S.). Colony-stimulating factor-granulocytes/macrophages (CSF-GM) was from Immunex, Seattle, Washington. CSF-granulocytes (CSF-G) was from Amgen, Thousand Oaks, California.

from macrophages when added directly to the scopoletin assay.

To determine whether the state of macrophage activation reflects the balance between activating and deactivating factors, macrophages were incubated in TGF- β_1 alone, macrophage activating factors alone, or combinations of these cytokines (Fig. 3). Suppression of H_2O_2 releasing capacity caused by TGF- β_1 could be overcome by interferon- γ , tumour necrosis factor (TNF)- α , or TNF- β , at concentrations similar to those required to activate resident peritoneal macrophages^{2,9,10}.

Suppression of H_2O_2 releasing capacity by TGF- β did not reflect toxicity to the macrophages. Thus, $\geq 95\%$ of macrophages excluded trypan blue after incubation for two days in $\geq 1\text{-ng ml}^{-1}$ TGF- β . Macrophages incubated for two days in 1- or 10-ng ml^{-1} TGF- β retained a capacity for extensive phagocytosis: treated macrophages took up ~ 20 –25 starch granules per cell (each granule 2.6 μm in diameter), marginally less than control cells, as assessed by a method optimized to detect differences in maximal phagocytic capacity¹¹ (Fig. 4). But TGF- β did cause macrophages to become less tightly adherent to the plastic culture surface. This was reflected by a dose-dependent decrease in the amount of cell protein remaining after vigorous washing ($\text{IC}_{50} \sim 0.8\text{ pM}$ for TGF- β_1 , $\sim 5.2\text{ pM}$ for TGF- β_2). Overall,

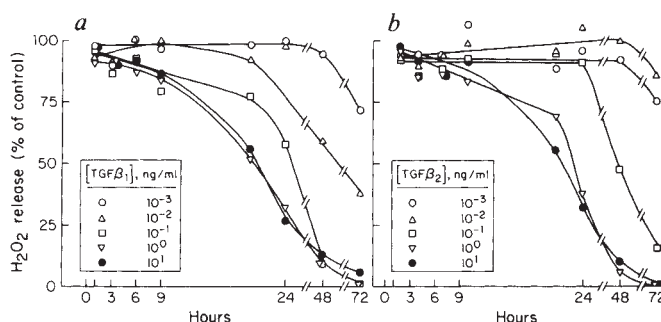


Fig. 2 Time course for suppression of H_2O_2 releasing capacity by (a) TGF- β_1 or (b) TGF- β_2 . After incubation for the periods indicated in 0.001–10 ng ml^{-1} TGF- β , the macrophages were washed and challenged with 167-nM PMA to measure H_2O_2 release (nmol per mg protein per 90 min). Values are expressed as a per cent of the results for controls incubated for the same time periods without TGF- β . The control values were nearly constant over three days. The mean $\pm$ s.e.m. for the means of eight control triplicates was 372 ± 13 in a and 362 ± 19 in b. One of two similar experiments is shown.

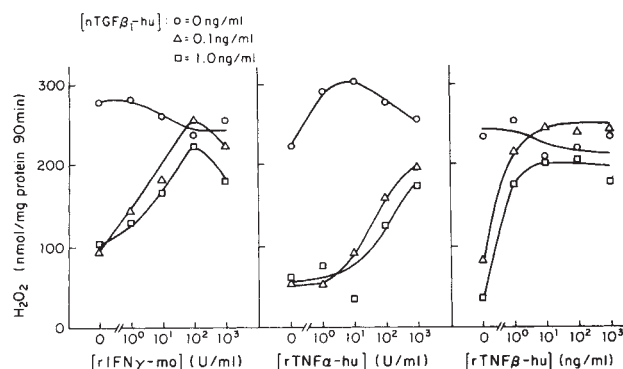


Fig. 3 Prevention of TGF- β_1 -induced deactivation by coinubation in macrophage activating factors. Activated macrophages were exposed to the following cytokines alone or in the indicated combinations: TGF- β_1 , and one of three pure recombinant proteins (all expressed in *Escherichia coli* and provided by Genentech)—a, murine interferon- γ (rIFN γ -mo, specific activity $5 \times 10^7\text{ U mg}^{-1}$), b, tumour necrosis factor- α -hu (TNF α , specific activity $3.6 \times 10^7\text{ U mg}^{-1}$) and (c) TNF- β -hu (lymphotoxin, specific activity $1.2 \times 10^8\text{ U mg}^{-1}$). After two days, the cells were washed and challenged with PMA. TGF- β_1 concentrations were 0 ($\circ$), 0.1 (Δ), or 1 ($\square$) ng ml^{-1} . Concentrations of the other cytokines are indicated on the abscissa. One of four similar experiments is shown.

after treatment for two days with 0.04–100 ng ml^{-1} TGF- β_1 , the adherent cell protein in vigorously washed monolayers averaged $59 \pm 20\%$ of control ($n = 83$ in 25 experiments).

Prominent sources of TGF- β include degranulating platelets, endothelial cells, fibroblasts, keratinocytes, tumour cells and T cells responding to antigen^{12–16}. Thus, wounds, tumours¹⁷ and T cells may all be able to restrain or reverse macrophage activation through the action of TGF- β . Eukaryotic cells can deactivate macrophages by other routes as well^{3–5}. But the previously described macrophage deactivating factor from tumours^{3–5} is immunochemically and functionally distinct from TGF- β (manuscript in preparation). Macrophages themselves can secrete TGF- β upon exposure to bacterial lipopolysaccharide¹⁸. This suggests the possibility of a negative feedback that may counter the potential for positive feedback implicit in the ability of the macrophage to produce some of its own activating factors, such as TNF- α ^{9,10} and CSF-GM¹⁹.

Other reported actions of TGF- β_1 on mononuclear phagocytes are both stimulatory and suppressive. TGF- β_1

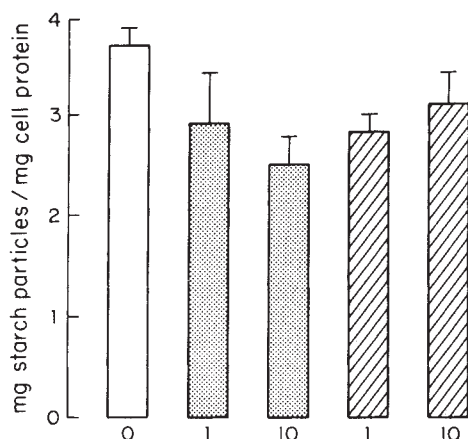


Fig. 4 Preservation of phagocytic function after deactivation of macrophages by TGF- β 1 or TGF- β 2. Macrophages (8.5×10^5) were plated on 13-mm glass coverslips and nonadherent cells removed after 2 h. Test media were added for 2 days as indicated: complete medium alone (open bar), or complete medium containing TGF- β 1 (stippled bars) or TGF- β 2 (hatched bars) at the concentrations indicated in ng ml^{-1} . After 2 days, macrophages were washed and incubated in complete medium for 2 h with 2 mg per coverslip of ^{14}C -acetylated starch granules isolated from seeds of *Amaranthus caudatus*. These conditions optimize the quantification of maximal phagocytic capacity by macrophages¹¹. The coverslips were washed and the monolayers solubilized to determine the mg particles phagocytized per mg cell protein. Means $\pm$ s.e.m. of triplicates are shown.

induces chemotaxis^{20,21} (EO_{50} , 0.004 pM)²⁰, release of fibroblast growth factors and accumulation of IL-1 mRNA (EC_{50} , $\sim 40 \text{ pM}$)²⁰, and release of angiogenic factors²¹. On the other hand, exposure of macrophages to TGF- β 1 is associated with $\sim 50\%$ suppression of TNF α release and $\sim 35\%$ suppression of Ia antigen expression (EC_{50} , $40\text{--}400 \text{ pM}$) (ref. 21 and personal communication, C. Czarnecki). Above, we have described virtually complete suppression of macrophage respiratory burst capacity by TGF- β 1 (EC_{50} , 0.6 pM) and TGF- β 2 (EC_{50} , 4.8 pM). These effects may reflect a coordinated response in wound healing²³, in which macrophages are recruited to scavenge debris and foster the growth of fibroblasts and endothelial cells, while being suppressed in their capacity for a respiratory burst that could be inimical to these cells.

The ability of TGF- β 1 and TGF- β 2 to ablate the respiratory burst of macrophages raises the possibility that these agents might have a role in the treatment of inflammatory disorders involving excessive macrophage activation.

We thank Moses Chao, Joshua Adler, Yuko Arita, Genentech, Inc., Amgen, Inc. and Immunex, Corp. for providing some of the cytokines. This work was supported by the National Cancer Institute.

Received 6 February; accepted 19 May 1988.

1. Confer, D. L. & Eaton, J. W. *Science* **217**, 948–950 (1982).
2. Ding, A. H. & Nathan, C. F. *J. Immunol.* **139**, 1971–1977 (1987).
3. Szuro-Sudol, A. & Nathan, C. F. *J. exp. Med.* **156**, 945–961 (1982).
4. Szuro-Sudol, A., Murray, H. W. & Nathan, C. F. *J. Immunol.* **131**, 384–387 (1983).
5. Tsunawaki, S. & Nathan, C. F. *J. exp. Med.* **164**, 1319–1331 (1986).
6. Cross, C. E. *et al. Ann. int. Med.* **107**, 526–545 (1987).
7. Nathan, C. F. *Trans. R. Soc. trop. Med. Hyg.* **77**, 620–630 (1983).
8. de la Harpe, J. & Nathan, C. F. *J. immun. Meth.* **78**, 323–325 (1985).
9. Murray, H. W., Spitalny, G. & Nathan, C. F. *J. Immunol.* **134**, 1619–1622 (1985).
10. Shparber, M. & Nathan, C. F. *Blood* **68**, 86a (1986).
11. Nathan, C. F. & Terry, W. D. *Cell. Immun.* **29**, 295–311 (1977).
12. Derynck, R. *et al. Nature* **316**, 701–705 (1985).
13. Kehrl, J. H. *et al. J. exp. Med.* **163**, 1037–1050 (1986).
14. Sporn, M. B., Roberts, A. B., Wakefield, L. M. & Assoian, R. K. *Science* **233**, 532–534 (1986).
15. Massague, J. *Cell* **49**, 437–438 (1987).
16. Akhurst, R. J., Fee, F. & Balmain, A. *Nature* **331**, 363–365 (1988).

17. Dvorak, H. F. *New Engl. J. Med.* **315**, 1650–1659 (1986).
18. Assoian, R. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 6020–6024 (1987).
19. Reed, S. G. *et al. J. exp. Med.* **166**, 1734–1746 (1987).
20. Wahl, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 5788–5792 (1987).
21. Wiseman, D. M., Polverini, P. J., Kamp, D. W. & Leibovich, S. J. *J. cell Biol.* **105**, 163a (1987).
22. Espevik, T. *et al. J. exp. Med.* **166**, 571–576 (1987).
23. Roberts, A. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 4167–4171 (1986).
24. Assoian, R. K., Komoriya, A., Meyers, C. A., Miller, D. A. & Sporn, M. B. *J. biol. Chem.* **258**, 7155–7160 (1983).
25. Marquardt, H., Lioubin, M. N. & Ikeda, T. *J. biol. Chem.* **262**, 12127–12131 (1987).

A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein

Jordi Gómez, Demetrio Sánchez-Martínez*, Virginia Stiefel, Joan Rigau, Pere Puigdomènech & Montserrat Pagès†

Departamento de Genética Molecular, CID (CSIC), Jorge Girona Salgado 18–26, 08034 Barcelona, Spain

Plant hormones such as abscisic acid (ABA) appear to modulate the responses of plants under adverse conditions^{1,2}. ABA has a poorly-understood role in embryogenesis, accumulating in the stages before dessication^{3,4}, and altering the rate of transcription of a specific set of genes^{5,6}. The functions of the proteins encoded by these genes, however, are unknown, and their messenger RNAs decrease again during early germination^{7–9}. No correlation has been established between ABA levels and the induction of particular genes in non-embryonic organs. The level of ABA increases substantially in leaf tissues subjected to water stress¹⁰ and thus it has been proposed that ABA mediates plant–water relations^{1,10}. Here we describe the isolation of complementary DNA and genomic clones of a gene that is ABA-inducible in the maize embryo, and whose messenger RNA accumulates in epidermal cells, which is also induced by water stress and wounding in leaves. The deduced protein is rich in glycine. Identification of this gene will contribute to our understanding of the role of ABA.

After a series of differential screenings of a cDNA library constructed from maize dry embryo using cDNA synthesized from immature embryo poly(A)⁺ RNA, with or without ABA treatment, six non cross-hybridizing clones were selected. These were shown to correspond to mRNAs present in dry embryos by Northern analysis, and their level increased precociously after ABA treatment of immature maize embryos. One of them (clone pMAH9) had an insert of 732 base pairs (bp) and hybridized to an RNA band of approximately the same size (Fig. 3). This clone was chosen for further study, and was used to screen a genomic library. The identity of a hybridizing genomic clone was confirmed by restriction mapping.

The sequences of the cDNA and genomic clones are shown in Fig. 1. They show perfect identity except for an insertion of 146 bp in the genomic sequence with the sequence features of an intron. There is a TATA box at the 5'-end. Only one plausible open reading frame was identified in the cDNA, and the sequence of the putative protein is shown in Fig. 1. The encoded protein is 157 amino acids long with a predicted relative molecular mass of 15,427 and an isoelectric point of 5.7. These values correspond closely with polypeptides detected in hybrid-released experiments using the pMAH9 plasmid (results not shown).

The protein sequence has a well-defined domain structure, clearly seen in hydrophilicity plots (Fig. 2). The first half of the sequence (residues 1–88) is composed of alternating α and short

* Present address: Herpes Branch, Centers for Disease Control, Atlanta, Georgia 30333, USA
† To whom correspondence should be addressed

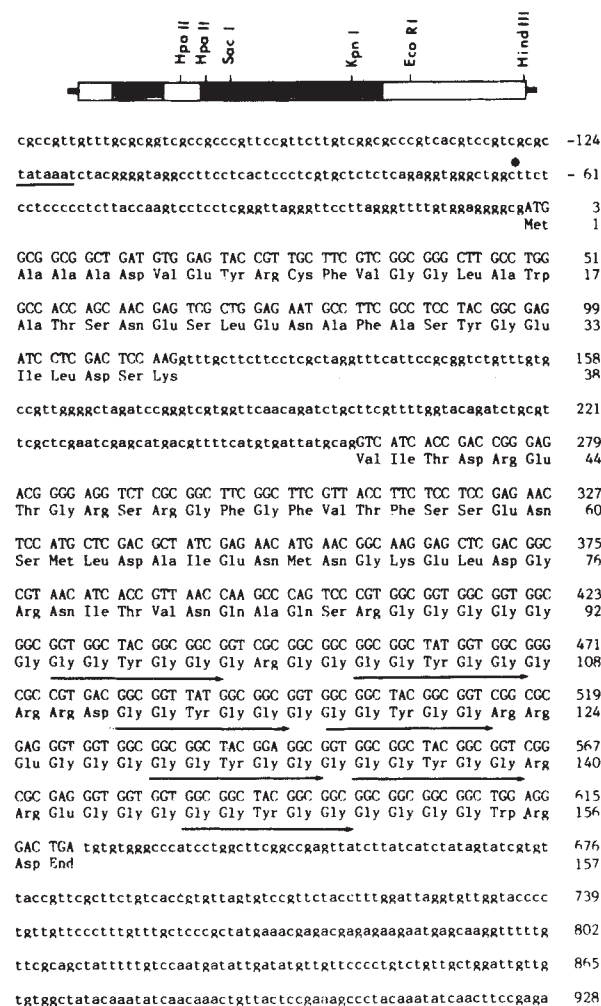


Fig. 1 Restriction endonuclease map and sequence of the insert of cDNA clone pMAH9 and the corresponding genomic clone. The sequenced zone is represented by a rectangle including the cDNA (hatched) and the main restriction sites. The nucleotide sequence includes a TATA box (underlined) and the limits of cDNA (shown by an asterisk). The open reading frame present in the sequence is shown (the repetitive Gly-Gly-Tyr-Gly-Gly motif is underlined with arrows). Clone pMH9 was selected by differential screening of a dry embryo cDNA library in pBR322 using single-stranded labelled cDNA probes synthesized from poly(A)⁺ RNA extracted from immature embryos (20 days after pollination) before and after treatment with 1 μ M ABA for 21 hours. The positive clones accounted for 2% of the cDNA library. The genomic clone was detected in a library constructed in λ charon35 after partial restriction of endosperm DNA with *Sau*III¹⁶. DNA sequence was carried out by chemical degradation¹⁷. cDNA and genomic sequences were identical except for the insert in the genomic fragment, containing the consensus sequences of an intron.

β structures, whereas the second half (residues 89–157) has an unusual amino-acid composition (Gly 69.4%, Arg 12.5%, Tyr, 9.7%, Asp 2.8%, Glu 2.8% and Trp 1.4%). The sequence of the second domain is highly repetitive with two main motifs, Gly-Gly-Tyr-Gly-Gly and Arg-Arg-Glu. The most abundant motif (Gly-Gly-Tyr-Gly-Gly) is different from that found in a putative cell-wall protein, rich in glycine, isolated from *putunia*¹¹.

Northern analysis and *in situ* hybridization were used to check the relative abundance of RNAs hybridizing with pMAH9. Hybridization of the cDNA insert with RNAs extracted from embryos and plants at different stages of development (Fig. 3) shows that a basal level of this mRNA exists in young roots

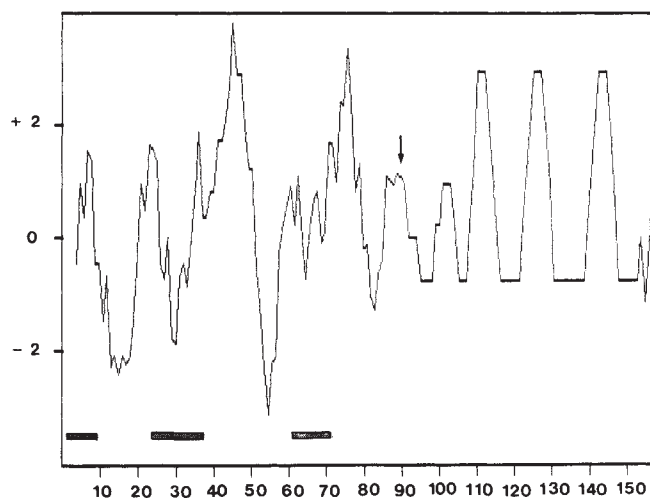


Fig. 2 Hydrophilicity plot of the predicted amino-acid sequence of the glycine-rich protein¹⁸. Hydrophilicity is plotted as a function of the residue number. The arrow indicates amino acid number 88, separating the two domain structures of the protein. The bars show the segments where α -structure is predicted¹⁹.

and leaves, and in immature embryos 20 and 30 days after pollination. This starts to increase 33–35 days after pollination (not shown) and the highest level is detected at 40 days, when embryo desiccation starts. The mRNA is also present in dry embryo but is barely detectable after 24 hours of germination. When young embryos are incubated in the presence of ABA a 25-fold increase in the level of this mRNA is seen.

In situ hybridization was carried out on cryostat sections of embryos 40 days after pollination to check whether the mRNA accumulates in a specific cell type (Fig. 4). The radioactivity accumulates in scutellar epidermal cells surrounding the embryo axis, and extends one layer of cells inwards. At higher magnification, strips of radioactivity are also seen in epidermal cells on the edges of the embryonic leaves forming the plumule (not shown). Several glycine-rich proteins^{12,13} which show limited tissue-specific expression have been described as having structural and protective functions. Whether the protein encoded by the gene described here has a similar function remains to be seen.

It has been proposed that endogenous ABA levels rise as a consequence of water stress in maize leaves and return to a basal level after rewatering¹⁴. We therefore examined whether pMAH9 mRNA is accumulated in leaves subjected to this and to other types of stress. Northern analysis of dehydrated leaves (Fig. 3) shows a high increase in the mRNA level, decaying to a basal level after recovery. During dehydration the ABA hormone concentration increased linearly to 10-fold over a three-hour period, and to maximum correlated with the maximum mRNA abundance. After rewatering, a close correlation between ABA concentration and mRNA levels was also observed. The effect of leaf wounding was also examined. The time course of mRNA accumulation after wounding in both young (not shown) and 30-day-old leaves (Fig. 3) shows that the maximum mRNA level is achieved after two hours; mRNA is still detectable at relatively high levels after 12 hours but decreases to basal levels by 24 hours.

These results indicate that the gene corresponding to pMAH9, selected as being ABA-inducible in maize embryo, is also induced in different situations when the corresponding organ is challenged by a hydric stress. This would occur in embryos as a prelude to the dormancy period, and is one of the primary effects of wounding upon leaf cells. We suggest that the increase in ABA levels in embryos before desiccation is part of the programme ensuring the survival of embryos during dormancy. The mechanisms of protection of cereals against hydric stress conditions are not well understood. Further characterization of

Fig. 3 Northern blot analysis of RNA hybridized with pMAH9. Total RNA was isolated from: *a*, excised embryos 20 days after pollination incubated in water for 21 h (lane 1) or in the presence of 1 μ M ABA (lane 2); freshly excised embryos 20 (lane 3), 30 (lane 4) and 40 (lane 5) days after pollination; dry embryos (lane 6); one-day germinated embryos (lane 7) and 7-day-old roots (lane 8). *b*, Young (7-day-old) leaves (lane 1), dehydrated leaves (lane 2) and rewatered leaves (lane 3). *c*, Wounded 30-day-old leaves 2 (lane 1), 12 (lane 2) and 24 (lane 3) hours after wounding; unwounded are shown as a control (lane 4). *d*, Dehydration of young leaves for 1, 2 and 3 h (D), and rehydration for 12 and 24 (R) control leaves (C). Numbers indicate the ABA content (pmol/g fresh weight) during and after dehydration. The blot was carried out with 10 μ g of RNA as previously described²⁰. The size of hybridizing mRNA is 800 nucleotides. Plants were dehydrated until they had lost 10–15% fresh weight by continuous exposure to a gentle stream of air during growing, and well watered plants were transferred to dry plates where they were dehydrated under a stream of air for 1, 2, and 3 hours. Rewatering of dehydrated plants was done during 12 and 24 hours. Leaves from 7 and 30 day-old plants were wounded by blade incision. Preparation of plant extracts for ABA measurements was done as described elsewhere²¹. The concentration of ABA was measured using a monoclonal antibody and ELISAs based on competitive binding between free and enzyme-linked ABA (Phytodetek-ABA kit from Idetek Inc., San Bruno, California).

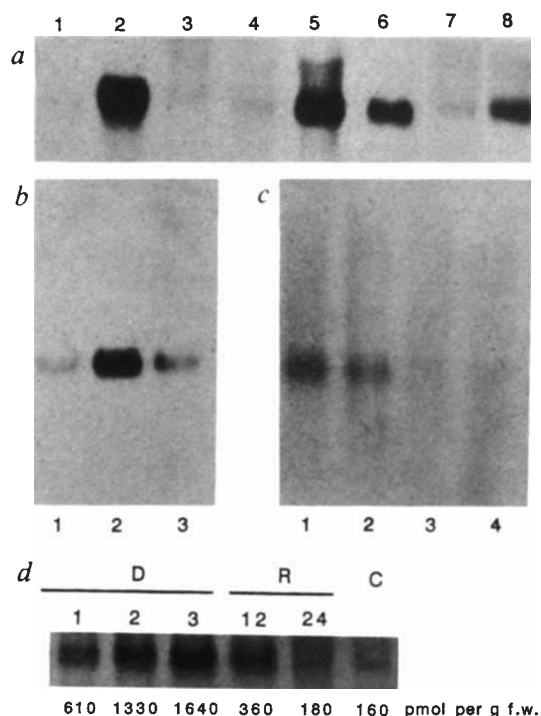
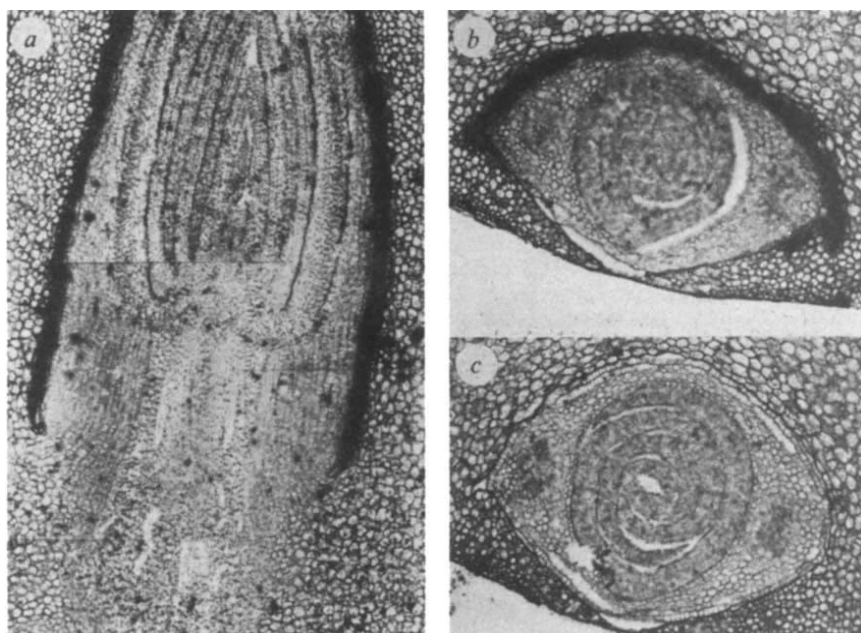


Fig. 4 Localization of mRNA complementary to pMAH9 cDNA on embryo (40 days after pollination) sections. The presence of the mRNA is observed in the scutellar epidermic cell layers surrounding the embryo axis. *a*, Longitudinal section of plumule-axis, *b*, transversal section of plumule and *c*, control transversal section of plumule hybridized with zein cDNA (clone A20²¹). Cryostat sections (8 μ m) of embryos of maize (W64A) were treated as previously described²². The cDNA insert was digested with KpnI, SacI and HpaII and labeled with ³⁵S nucleotides by random priming (Boehringer Mannheim kit). Hybridization and washings were performed as described but DTT (10 mM) was added in all buffers. The specific activity of the probe was 1×10^8 cpm/ μ g and it was used at 10–30 ng per slide. Autoradiography was done with Kodak NTB-2 emulsion. Replicate sections were preincubated with ribonuclease A (100 μ g/ml) and ribonuclease T1 (5 μ g/ml) as a control for non-specific binding.



the gene described here should provide significant insights into the role of ABA in plant stress tolerance.

We thank Dr S. Prat for practical suggestions, Drs F. and B. Burr for providing clone A20, Ms M. Alvarez for manuscript preparation, and especially Dr M. D. Ludevid for performing the ABA determinations. This work was supported by grants from CSIC and CSIC-CDTI (Plan Proa).

Note added in proof: A recent report of Chandler *et al.*²⁴ confirms our results, showing the expression of ABA-inducible genes in water stressed cereal seedlings.

Received 24 March; accepted 10 June 1988.

1. Davies, M. J. & Mansfield, T. A. in *Abscisic Acid* (ed. Addicott, F. T.) 237–268 (Praeger, New York, 1983).
2. Singh, N. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 739–743 (1987).
3. King, R. W. *Planta* **132**, 43–51 (1976).
4. Jones, R. J. & Brenner, M. L. *Pl. Physiol.* **83**, 905–909 (1987).

5. Williamson, J. D., Quatrano, R. S. & Cuming, C. *Eur. J. Biochem.* **152**, 501–507 (1985).
6. Baker, J. C. & Dure, L. in *Molecular Biology of Plant Growth Control* (eds Fox, J. E. & Jacobs, M.) 51–62 (Liss, New York, 1987).
7. Galau, G. A., Hughes, D. W. & Dure, L. *Plant molec. Biol.* **7**, 155–170 (1986).
8. Quatrano, R. S. *Oxford Surv. Plant molec. Cell Biol.* **3**, 467–477 (1986).
9. Sánchez-Martínez, D., Puigdoménech, P. & Pagés, M. *Pl. Physiol.* **82**, 543–549 (1986).
10. Wright, S. T. C. & Hiron, R. W. P. *Nature* **224**, 719–720 (1969).
11. Condit, C. H. & Meagher, R. B. *Nature* **323**, 178–181 (1986).
12. Hannkoglin, I. & Fuchs, E. *Cell* **33**, 915–924 (1983).
13. Eickbush, T. H. & Burke, W. D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2814–2818 (1985).
14. Beardsell, M. F. & Cohen, D. *Pl. Physiol.* **56**, 207–212 (1975).
15. Pauling, L. & Corev, R. B. *Proc. natn. Acad. Sci. U.S.A.* **37**, 729–740 (1951).
16. Gallardo, D. *et al. Pl. Sci.* **54**, 211–218 (1988).
17. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
18. Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824–3828 (1981).
19. Garnier, J., Osguthorpe, D. J. & Robson, B. *J. molec. Biol.* **120**, 97–117 (1978).
20. Pérez-Grau, L. *et al. FEBS Lett.* **202**, 145–148 (1986).
21. Raikhel, N. V., Hughes, D. W. & Galau, G. A. in *Molecular Biology of Plant Growth Control* (eds Fox, J. E. & Jacobs, M.) 197–207 (Liss, New York, 1987).
22. Burr, B. *et al. J. molec. Biol.* **154**, 33–49 (1982).
23. Sánchez-Martínez, D. *et al. Pl. Sci.* **53**, 215–222 (1987).
24. Chandler, P. M. *et al. J. Cell. Biol. Suppl.* **12C** Abs L-033, 143 (1988).

Self-cleaving viroid and newt RNAs may only be active as dimers

Anthony C. Forster*, Christopher Davies,
Candice C. Sheldon, Alex C. Jeffries
& Robert H. Symons†

Department of Biochemistry, University of Adelaide, Adelaide,
South Australia 5000, Australia

Avocado sunblotch viroid (ASBV) is a 247-nucleotide, single-stranded, circular RNA¹. It is considered to replicate via a rolling-circle mechanism²⁻⁴ in which circular, monomeric plus and minus RNAs act as templates for the synthesis of longer-than-unit-length precursor RNAs. Processing of these RNAs *in vivo* may occur by a self-cleavage reaction, as indicated by ability of dimeric, linear plus and minus ASBV RNAs to specifically self-cleave *in vitro* with the excision of a monomeric RNA with 5'-hydroxyl and 2',3'-cyclic phosphodiester termini⁴. A similar self-cleavage reaction has also been reported to occur in an RNA transcript containing a dimeric copy of a tandemly repeated, 330-base-pair sequence of the newt genome⁵. Based on comparisons with self-cleaving plant viral satellite RNAs^{6,7}, hammerhead-shaped active structures, each containing one self-cleavage site, were proposed for the plus and minus ASBV RNAs⁴ and the newt RNA⁵, but the stability of these hammerheads has been questioned^{4,8}. Here, more stable active structures that contain two self-cleavage sites are proposed and data supporting these models are presented.

Figure 1a shows the previously proposed hammerhead models for the self-cleavage structures of the ASBV⁴ and newt⁵ RNAs, and Fig. 1b shows the consensus hammerhead structure found at the known or predicted self-cleavage sites of seven plant viral satellite RNAs (including six virusoid RNAs)⁶. Previous work has shown that the hammerhead sequences of the virusoid of lucerne transient streak virus (vLTSV), minus ASBV and newt RNAs are necessary and sufficient for cleavage⁸⁻¹¹. In contrast to the theoretically stable stem IIIs of the satellite RNAs (Fig. 1b), the ASBV stem IIIs (Fig. 1a) are theoretically unstable^{4,8}. Furthermore, the newt RNA stem III (Fig. 1a), which has only a two-base hairpin loop, is unlikely to form due to steric hindrance¹².

* Present address: Department of Biology, Kline Biology Tower, Yale University, New Haven, Connecticut 06520, USA.

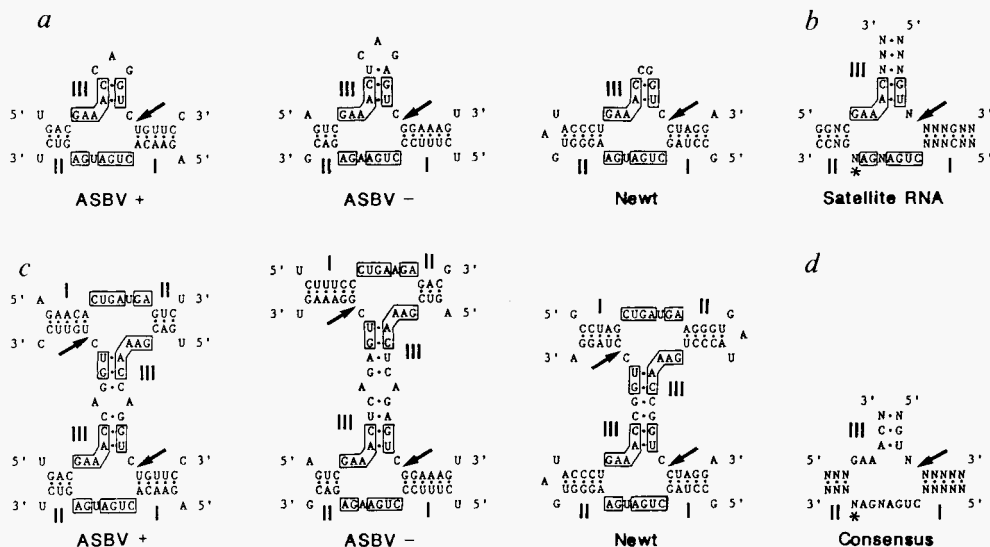
† To whom correspondence should be addressed.

Examination of the sequences of the self-cleaving RNAs shows that it is possible to construct secondary structures containing two hammerheads. In these structures, both the plus and minus ASBV RNAs and the newt RNA contain stem IIIs that are theoretically stable¹³. Figure 1c (refer also to Fig. 4b) shows the double-hammerhead-shaped secondary structures formed by the association of two plus ASBV, minus ASBV or newt hammerheads. Deletion analysis of the plus sense vLTSV RNA has shown that only the two lowermost base pairs of stem III (see Fig. 1b) are required for self-cleavage⁸. If this criterion is extended to the double-hammerhead structure (Fig. 1c) it would seem that either the top or bottom halves of these structures alone are sufficient for cleavage at a single site. Thus, a consensus hammerhead structure (Fig. 1d) can be drawn for all the RNAs in Fig. 1b, c.

To distinguish between active structures containing one hammerhead and those containing more than one hammerhead, a 40-mer RNA containing the sequence of a single newt-like hammerhead was synthesized from the appropriate DNA template by T7 RNA polymerase¹⁴ and the relationship between the extent (%) of self-cleavage and RNA concentration determined. Figure 2a shows a pair of these RNAs in the double-hammerhead arrangement. The promoter requirements for the T7 RNA polymerase system¹⁴ did not allow us to prepare an RNA identical in sequence to that derived from the newt RNA (Fig. 1a). This newt-like RNA cleaved specifically and direct RNA sequence analysis¹⁵ established that cleavage had occurred at the expected bond (indicated by arrows in Fig. 2a). Examination of the extent (%) of self-cleavage of the newt-like RNA over a 1,000-fold concentration range revealed that it increased with increasing concentration (Fig. 2b). The initial rate of self-cleavage ($\text{ng } \mu\text{l}^{-1} \text{ min}^{-1}$) at a concentration of $50 \text{ ng } \mu\text{l}^{-1}$ of RNA was ~80 times that at an RNA concentration of $5 \text{ ng } \mu\text{l}^{-1}$; the second-order rate equation for a bimolecular reaction predicts a 100-fold difference in rate. This result indicates that cleavage occurred by the association of two molecules, as predicted by the double-hammerhead model.

In contrast, the isolated 52-nucleotide vLTSV hammerhead sequence, containing the minimum sequence for self-cleavage, was cleaved completely in less than one minute at 25 °C at a concentration of $0.2 \text{ pg } \mu\text{l}^{-1}$ (in 50-mM Tris-HCl, pH 7.5, 5-mM MgCl_2 , 0.5-mM EDTA)⁸. Under similar conditions the newt-like 40-mer RNA containing the hammerhead sequence (Fig. 2a) cleaved to a negligible extent (data not shown). This can be explained by considering the vLTSV hammerhead reaction as intramolecular (as the hammerhead structure contains a stable

Fig. 1 Hammerhead structures of self-cleaving viroid, newt and satellite RNAs. **a**, Previously published hammerhead structures for plus and minus ASBV RNAs⁴ and newt RNA⁵. **b**, The consensus sequence for the hammerhead structures found in seven plant satellite RNAs⁶ (that is, plus and minus virusoid RNAs of two isolates of lucerne transient streak virus (vLTSV-A and -N), the plus virusoid RNAs of velvet tobacco mottle virus and solanum nodiflorum mottle virus and the plus satellite RNA of tobacco ringspot virus). **c**, The proposed double-hammerhead structures for plus and minus ASBV RNAs and newt RNA. **d**, Consensus hammerhead structure for ASBV, newt and plant satellite RNAs. Stems are numbered I to III (after convention of Forster and Symons⁶), sites of cleavage are indicated by arrows and nucleotides conserved between ASBV, newt and plant satellite RNAs (see Fig. 1b legend) are boxed. The asterisked nucleotide is unique to the sequence of the vLTSV plus hammerhead where it is a U.



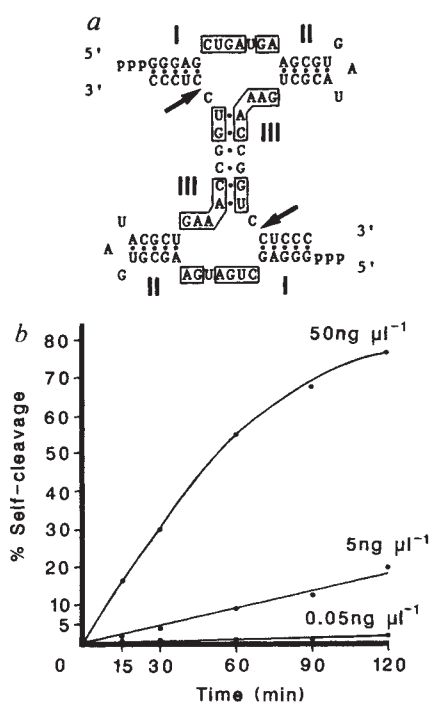


Fig. 2 Effect of concentration of newt-like RNA on the extent (%) of self-cleavage. *a*, Double-hammerhead secondary structure for the newt-like 40-mer RNA generated by T7 RNA polymerase transcription¹⁴ of a synthetic DNA template. Labelling of stems, sites of cleavage and boxing of conserved regions as in Fig. 1. *b*, Plot of per cent self-cleavage of newt-like 40-mer RNA at three different concentrations (0.05, 5 and 50 ng μl^{-1}) versus time.

Method. Non-radioactive and ^{32}P -labelled RNAs were produced by transcription of a synthetic DNA template by T7 RNA polymerase¹⁴ under essentially the same conditions as in Forster and Symons⁸. Full-length transcripts were purified by polyacrylamide gel electrophoresis, eluted and ethanol precipitated⁶. RNA concentrations were determined by spectroscopy for non-radioactive RNA and by liquid scintillation counting for ^{32}P -labelled RNA. 0.05 ng μl^{-1} of ^{32}P -labelled RNA was used in all reactions and non-radioactive RNA added to achieve the required RNA concentrations. RNAs in 1 mM EDTA, pH 6.0, were heated at 80 °C for 1 min, snap-cooled on ice and then incubated in 50 mM Tris-HCl, pH 8.0, 10 mM MgCl_2 , 0.5 mM EDTA (5 μl final volume) at 55 °C for various times. Reactions were terminated by the addition of EDTA to 10 mM and an equal volume of formamide⁶. Products were resolved on a 7 M urea, 20% polyacrylamide gel and identified by autoradiography. The bands were excised and liquid scintillation counting used to determine the extent of self-cleavage.

stem III, Fig. 1*b*), and the newt hammerhead reaction as intermolecular (because stable stem IIIs may only be formed by the association of two RNA molecules). Indeed all the satellite RNA hammerheads, which contain stable stem IIIs (Fig. 1*b*), are likely to undergo intramolecular self-cleavage reactions.

To demonstrate the requirement for a double-hammerhead in plus ASBV RNA self-cleavage an approach involving dimeric RNA transcripts was used. This system is more likely to resemble the *in vivo* situation^{2,4} than short oligonucleotides as sequential production and folding of the full-length RNA can occur. Dimeric plus ASBV RNA transcripts (5'E/M/3'E) prepared *in vitro* from a dimeric complementary DNA template⁴ cleaved specifically and efficiently during transcription to yield a monomeric RNA (M) of 247 nucleotides (Figs. 3*a, b*, lane 2; ref. 4). When the dimeric template was truncated by digestion with *Bcl*I (Fig. 3*a*), and transcribed with SP6 RNA polymerase a plus monomeric RNA (M*) was produced. This monomeric RNA underwent negligible cleavage during transcription (Fig.

3*b*, lane 1), even though it contained an intact self-cleavage sequence. This result supports the double-hammerhead model because the monomeric RNA molecule can only form an active self-cleavage structure (that is, a structure with stable stem IIIs, Fig. 1*c*) if two molecules become associated during the transcription reaction to form a double hammerhead. The formation of an intermolecular double hammerhead structure is expected to be much less rapid than formation of the equivalent structure within a single dimeric RNA molecule. The folded dimeric plus ASBV RNA, containing the proposed double hammerhead, is schematically represented in Fig. 4*b*; for comparison, dimeric ASBV RNA folded to contain two single-hammerhead structures is given in Fig. 4*c*.

Mutated dimeric plus ASBV RNAs produced by SP6 RNA polymerase transcription of site-directed mutagenized¹⁶ dimeric cDNA templates provided convincing evidence for the double-hammerhead model. In these experiments either or both of the GAAAC sequences (labelled A and B in Fig. 4*a, b, c*), conserved in all the RNAs given in Fig. 1, were mutated to GAAC. When both GAAACs were mutated, cleavage at both sites was abolished (Fig. 3*b*, lane 3), demonstrating the importance of this sequence for cleavage. When only the GAAAC nearest the 5'-end (labelled A in Fig. 4*a, b, c*) was mutated, cleavage at self-cleavage site one (SC-1 in Fig. 4*a, b, c*), which is spatially removed from the mutated GAAAC in the double-hammerhead model, was unaffected whereas cleavage at self-cleavage site two (SC-2) was abolished as indicated by the presence of only the 5'E and M/3'E fragments (Fig. 3*b*, lane 4). In the single-hammerhead model the reverse would be expected (compare with Fig. 4*b, c*). Similarly, mutation of the GAAAC nearest the 3' end (labelled B in Fig. 4*a, b, c*) abolished cleavage at SC-1 but did not affect cleavage at SC-2 as indicated by the presence of only the 5'E/M and 3'E fragments (Fig. 3*b* lane 5). Hence, these results provide strong evidence for the double-hammerhead model for plus ASBV. Investigation of minus ASBV self-cleavage is currently underway in our laboratory.

All available evidence indicates that ASBV replicates via a rolling-circle mechanism in which multimeric precursor RNAs self-cleave to yield linear monomeric plus and minus strands that are subsequently ligated to give the circular monomeric forms²⁻⁴. In such a replication scheme, efficient self-cleavage of longer-than-unit-length RNAs is necessary, but cleavage of the circularized monomeric RNAs is futile. The double-hammerhead model provides a mechanism whereby cleavage of monomeric ASBV RNAs is prevented but longer-than-unit-length RNAs containing two or more cleavage sites can cleave efficiently. This model also suggests that the release of excised ASBV monomer from longer-than-unit-length plus and minus ASBV RNAs occurs via the dissociation of the base-pairing in the two stem IIIs (Figs 1*c, 4b*). The ends of each monomer would then be juxtaposed for circularization by a plant RNA ligase(s).

The double-hammerhead model may have important implications for the enzyme/substrate system of Uhlenbeck¹¹ which used minus ASBV-like sequences whereby a 19-mer (01) RNA was shown to be capable of catalytically cleaving a 24-mer (02) RNA, and for the similar system of Koizumi *et al.*¹⁰ which used two 21-mer RNAs containing the newt self-cleavage sequences. These RNAs may undergo cleavage via a double-hammerhead structure formed by the association of four RNA molecules rather than by the previously proposed single-hammerhead structures^{10,11} composed of two RNA molecules with potentially unstable stem IIIs. It is also possible that an active self-cleavage structure containing a stable stem III may be formed as an incomplete double-hammerhead by the association of only one catalytic RNA and two substrate RNAs, only one of which would be cleaved.

The double-hammerhead model is especially relevant to future efforts to elucidate the tertiary structure of such self-cleaving molecules. Attempts to use short RNAs containing plus ASBV-

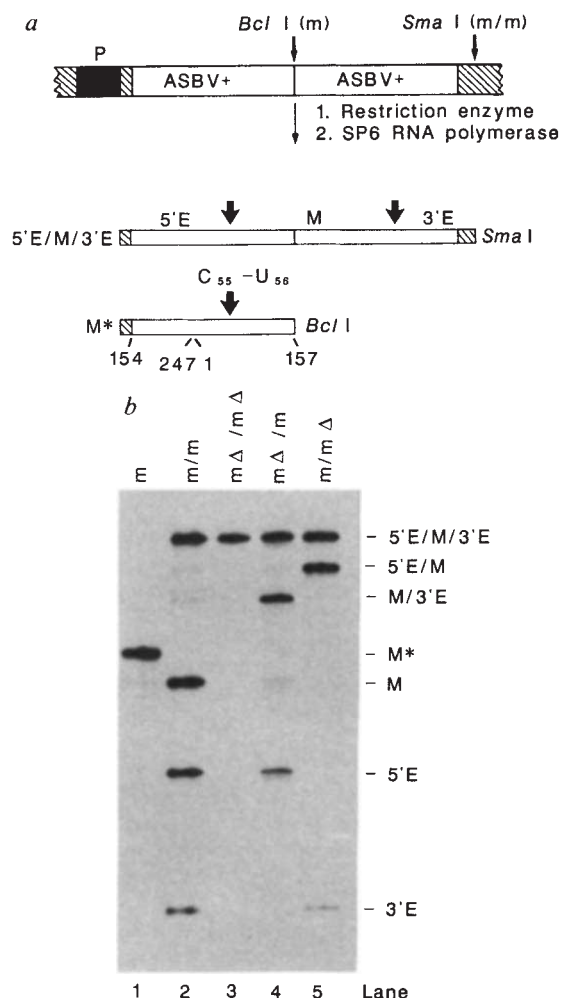


Fig. 3 Synthesis and cleavage of monomeric and dimeric plus ASBV RNAs; effect of GAAAC to GAAC mutation on ability of dimeric plus ASBV RNA to self-cleave. *a*, Diagram of plus ASBV dimeric DNA template and full-length RNAs produced by SP6 RNA polymerase transcription of templates linearized with either *Sma*I (m/m) or *Bcl*I (m). Striped regions represent vector sequences; closed box, SP6 promoter; open boxes, ASBV sequence. Sites of self-cleavage are indicated by thick arrows and ASBV sequence is numbered after Symons¹. Self-cleavage at both sites of the full-length transcript of *Sma*I linearized template (m/m) gives rise to a 5'-end fragment (5'E), a 247-nucleotide monomeric fragment (M) and a 3'-end fragment (3'E). *b*, Autoradiogram of polyacrylamide gel showing RNA transcripts of plus ASBV templates. Lane 1; transcript M* of wild-type dimeric ASBV template truncated with *Bcl*I (m, see Fig. 3a). Lane 2; as for lane 1 but template linearized with *Sma*I (m/m). Lane 3; as for lane 2 but both template GAAAC sequences (A and B) (see Figs 1 and 4a, b, c) mutated to GAAC (m Δ/mΔ). Lane 4; as for lane 3 but only GAAAC (A) mutated (m Δ/m). Lane 5; as for lane 3 but only GAAAC (B) mutated (m/m Δ). Labels down margin represent RNA fragments, either full-length (5'E/M/3'E) for *Sma*I digested template, M* for *Bcl*I digested template, or fragments resulting from self-cleavage (see Fig. 3a).

Method. Monomeric ASBV clones in phage M13mp93 (ref. 17) were mutated by oligonucleotide-directed mutagenesis essentially as described by Zoller and Smith¹⁶. The oligomer used was a 30-mer and resulted in mutation of the GAAAC sequence (see Figs 1a and 4a, b, c) to GAAC. Mutants were confirmed by dideoxy sequencing¹⁸. Mutant and wild-type monomeric inserts were excised, purified, mixed at a 1:1 ratio and ligated with T4 DNA ligase to form dimers. Dimers were purified and cloned into pSP64 (ref. 19). Identity of inserts was checked by subcloning into M13 followed by dideoxy sequencing. Selected clones were digested with either *Sma*I or *Bcl*I (Fig. 3a, b), transcribed with SP6 RNA polymerase (including α -³²P-UTP) and transcripts run on a 7 M urea, 5% polyacrylamide gel as described by Forster and Symons⁶.

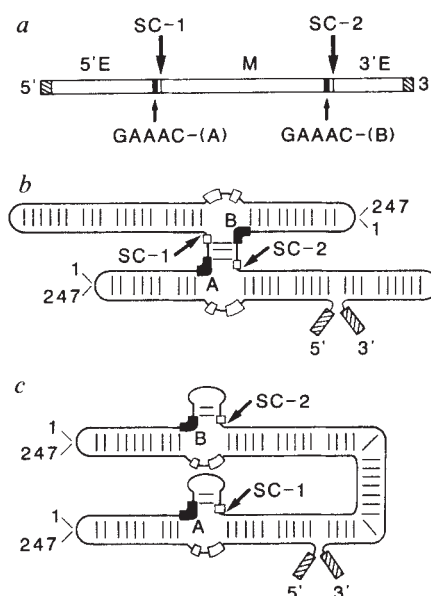


Fig. 4 Schematic representation of plus dimeric ASBV RNA transcript unfolded and folded in single- and double-hammerhead conformations. *a*, Unfolded SP6 RNA polymerase transcript of plus dimeric ASBV clone (see Fig. 3a). Self-cleavage sites, labelled SC-1 and SC-2, are indicated by arrows; striped boxes, vector sequences at 5'- and 3'-ends; closed boxes, GAAAC sequences (see Fig. 1) labelled A and B indicated by arrows. *b*, Schematic diagram of plus ASBV dimeric RNA drawn containing a double-hammerhead. Labelled as for Fig. 4a except for open boxes, conserved nucleotides (see Fig. 1); base-pairing is represented by lines between RNA strands. Sequence numbered after Symons¹. *c*, Schematic diagram of plus ASBV dimeric RNA transcript drawn to contain two single hammerheads. Labelled as in Fig. 4b.

like or newt-like sequences (that is, those lacking stable stem IIIs) for X-ray crystallographic or NMR studies may be complicated by the requirement for intermolecular interaction and the presence of two self-cleavage sites per unit structure. Better substrates for such experiments may be structures, such as the vLTSV hammerhead, that contain stable stem IIIs (Fig. 1b).

We thank Steven Rogers for the synthesis of the oligonucleotides, Paul Krieg, Sidney Altman, Fil Lim, Peter Rathjen and John Coates for discussions and comments on the manuscript and Jennifer Cassady and Tamara Greatrex for assistance. This work was supported by the Australian Research Grants Scheme and the Commonwealth Special Research Centre for Gene Technology.

Received 6 April; accepted 14 June 1988.

- Symons, R. H. *Nucleic Acids Res.* **9**, 6527-6537 (1981).
- Branch, A. D. & Robertson, H. D. *Science* **223**, 450-455 (1984).
- Hutchins, C. J. *et al. Pl. molec. Biol.* **4**, 293-304 (1985).
- Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. *Nucleic Acids Res.* **14**, 3627-3640 (1986).
- Epstein, L. M. & Gall, J. G. *Cell* **48**, 535-543 (1987).
- Forster, A. C. & Symons, R. H. *Cell* **49**, 211-220 (1987).
- Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. *Science* **231**, 1577-1580 (1986).
- Forster, A. C. & Symons, R. H. *Cell* **50**, 9-16 (1987).
- Forster, A. C., Jeffries, A. C., Sheldon, C. C. & Symons, R. H. *Cold Spring Harb. Symp. quant. Biol.* (in the press).
- Koizumi, M., Iwai, S. & Ohtsuka, E. *FEBS Lett.* **228**, 228-230 (1988).
- Uhlenbeck, O. C. *Nature* **328**, 596-600 (1987).
- Tinoco, I., Uhlenbeck, O. C. & Levine, M. D. *Nature* **230**, 362-367 (1971).
- Freier, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9373-9377 (1986).
- Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. *Nucleic Acids Res.* **15**, 8783-8798 (1987).
- Haseloff, J. & Symons, R. H. *Nucleic Acids Res.* **9**, 2741-2752 (1981).
- Zoller, M. J. & Smith, M. *Meth. Enzym.* **100**, 468-500 (1983).
- Barker, J. M., McInnes, J. L., Murphy, P. J. & Symons, R. H. *J. virol. Meth.* **10**, 87-98 (1985).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035-7055 (1984).

Molecular cloning and complete amino-acid sequence of form-I phosphoinositide-specific phospholipase C

C. Frank Bennett, Joanna M. Balcarek, Angela Varrichio, & Stanley T. Crooke

Department of Molecular Pharmacology,
Smith Kline & French Laboratories,
709 Swedeland Road, King of Prussia, Pennsylvania 19406, USA

We report the molecular cloning and sequence of a phosphoinositide-specific phospholipase C (PI-PLC), an enzyme that is of particular interest because of its central role in cell signal transduction. The signals in question are those delivered by hormones to their cell-surface receptors that activate PI-PLC by means of a guanine nucleotide binding protein. Activation of the enzyme leads to the hydrolysis of phosphatidylinositol 4,5-bisphosphate to two second messengers, 1,2-diacylglycerol and inositol 1,4,5-trisphosphate, the second of which ultimately mobilizes internal pools of calcium^{1,2}. There are at least five PI-PLC isoenzymes, whose differences in structure and function are unknown³⁻¹⁰. We have focused on isoenzyme I, which we have recently purified and characterized from guinea pig uterus⁸. We have now determined the sequence of a full length complementary DNA of this isoenzyme from the rat. Although the sequence has little similarity with the only other sequenced PI-PLC isoenzyme¹¹, it has a surprising degree of similarity to thioredoxins, protein co-factors in thiol-dependent redox reactions¹².

Antibodies prepared against guinea pig uterus PI-PLC I, previously demonstrated to react with a similar enzyme from rat basophilic leukaemia cells (RBL-1)⁸, were used to screen an RBL-1 λ gt11 cDNA expression library¹³. Initial screening of 200,000 plaques with the PI-PLC I antibodies yielded six immunoreactive clones, all of which were found to cross hybridize. The clone containing the longest insert, λ PLC-11 (Fig. 1) was used to affinity purify antibody from the original antiserum¹³. The resulting 'selected' antibody reacted specifically with PI-PLC I on immunoblots of total guinea pig uterus cytosolic proteins (data not shown), demonstrating the presence of PI-PLC I epitopes on the fusion protein. PLC-11 was also found to hybridize to an oligonucleotide probe generated from NH₂-terminal protein sequence of guinea pig uterus PI-PLC I.

DNA sequencing of PLC-11 (Fig. 1) revealed that the 1.8 kilobase (kb) clone encoded an open reading frame 1.5 kb in length and included 0.3 kb of 3' untranslated sequence through to the polyadenylation signal. The predicted protein sequence contained peptides of identical sequence to that obtained for five tryptic fragments derived from guinea pig uterus PI-PLC (Fig. 2). Moreover, PLC-11 also encoded the NH₂-terminal sequence of the guinea pig uterus protein with the exception of two amino-acid substitutions, Thr for Trp at position 31 and a Glu for Asp at position 33 (Fig. 2), changes which could reflect species differences or differences in isoenzymes. Although the open reading frame extended 42 base pairs (bp) 5' to the putative amino terminus, we were unable to identify an AUG initiation codon. Therefore, a second RBL-1 λ gt-11 library was screened with the 5' *Eco*R1-*Ava*1 fragment of PLC-11. Several longer clones were identified, including PLC-79, which was found to contain an additional 100 bp 5' to the start of PLC-11 (Fig. 1). DNA sequencing revealed that the two clones were identical and that the open reading frame of PLC-79 extended 30 bp further upstream than PLC-11 to an AUG initiation codon, encoding a protein of 504 amino acids with a relative molecular mass (M_r) of 56,559 (Fig. 2).

The observation that PI-PLC I was NH₂-terminally processed was unexpected. The 24 amino-acid sequence not present in the

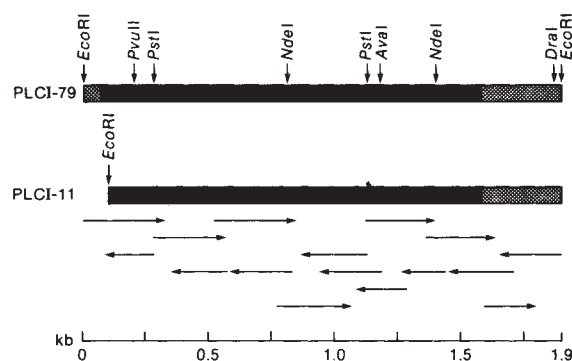


Fig. 1 Restriction enzyme map of PI-PLC I cDNA clones. PI-PLC I cDNA clones, PLC-79 and PLC-11, were restriction mapped and aligned as indicated. Both strands of PLC-11 were sequenced as indicated by the arrows using the dideoxy terminator method²¹ using a modified T7 DNA polymerase²² (Sequenase Kit, USB). Sequence was obtained by subcloning fragments of PLC-11 into m13 and by use of synthetic primers. Sequence of PLC-79, subcloned into m13, was obtained using synthetic primers. Overlapping sequence of PLC-11 and PLC-79 was determined to be identical.

purified cytosolic protein bears a strong resemblance to NH₂-terminal signal sequences¹⁴ and membrane anchoring domains which have been reported for many other proteins^{14,15}. This is particularly interesting given the controversial relationship between the cytosolic and membrane-associated forms of PI-PLC⁶⁻⁹. Although our initial experiments did not detect any differences between cytosolic and membrane-associated PI-PLC I, nor any processing of the enzyme⁸, we speculate that this region may be important in membrane interactions.

The predicted amino-acid sequence contains several potential serine/threonine phosphorylation sites (119-130, 225-234, 270-275), which supports our previous data that PI-PLC I is phosphorylated by protein kinase C *in vivo* and *in vitro*⁸. There is, however, no tyrosine phosphorylation consensus sequence and, in agreement with our biochemical characterization, no consensus sequence for N-linked glycosylation. We were unable to identify a E-F hand calcium binding site¹⁶ or a lipocortin-type calcium binding consensus sequence¹⁷. A database search revealed no striking similarities with other previously sequenced phospholipases (for example, PLA₂). Hydropathy analysis revealed that the NH₂-terminal 75 amino acids contain two hydrophobic domains, one of which is found within the first 24 amino acids. The remainder of the protein is overall rather hydrophilic, especially the COOH-terminus which contains five consecutive Lys residues. PI-PLC I also contains six repeats of three consecutive acidic amino-acid residues, one of which may be methylated.

A search for additional internal repeats revealed that PI-PLC I contained a 100 amino-acid sequence which is repeated twice within the protein sequence, at positions 24-133 and again at positions 374-484. The repeat was 57% similar overall with 34% amino-acid identity (Fig. 3). A search of the protein database (Dayhoff) revealed a high degree of similarity between the PI-PLC I repeat sequence and thioredoxins¹² (Fig. 3). Thioredoxins are small proteins which act as co-factors in a variety of biologically important oxidation-reduction reactions, including ribonucleotide reduction and photosynthesis¹². The active redox site of thioredoxin contains a cystine disulphide ring (Cys-Gly-Pro-Cys) comprised of 14 atoms, which is highly conserved for all the thioredoxins. Both PI-PLC I repeats also contain a 14-membered cystine disulphide ring which is identical to that found in thioredoxins except a His is substituted for Pro (Cys-Gly-His-Cys) (Fig. 3). Moreover, there is a high degree of similarity in the amino acids adjacent to the thioredoxin active

Fig. 2 Sequence of rat PI-PLC I cDNA and predicated amino-acid sequence. DNA sequence of PLC-79 was determined as described in Fig. 1. Nucleotides specifying the NH₂-terminal methionine, translation stop codon, and polyadenylation signal are underlined. Predicted amino-acid sequence corresponding to amino-acid sequence derived from protein sequence analysis of guinea pig uterus PI-PLC I tryptic peptides are also underlined. Protein sequence corresponding to NH₂ terminus of guinea pig uterus PI-PLC I purified from cytosolic fractions are also indicated (=). Protein sequencing was performed on a Beckman model 890 M sequencer.

```

1  GATTGGGCTCTCCCATCTCGGTTCTGTGGTCCCGCCCTCCGATTGCCCGCCGATCCCGCCGCCATGCTTCAGCTGCCTGCCTGCTCCCGG
MetProSerAlaAlaLeuArgCysSerArg 10
100 GCGTGGGATTCCTGCTCGCTCGCCCTCTCGCTCCGCTCAGACGTGTGGAACTGACGACGACAAATTCAGAGTCCGCTCTCCGACACGGG
AlaTrpArgLeuLeuLeuAlaSerAlaLeuLeuAlaSerAlaSerAspValLeuGluLeuThrAspGluSerArgThrGly 43
199 TCAGCTGGCTATGCTAGTCGAGTCTCTCCGCTATGGTGGACATTGCAAGAGGCTTGCCTGAGTATGAAGCTCAGCAACAGGATTAAGGA
SerAlaGlyLeuMetLeuValGluPhePheAlaProTrpCysGlyHisCysLysArgLeuAlaProGluGlyAlaAlaThrArgLeuLysGly 76
298 ATAGTCCCATAGCAAGGTGGACTGCACGCCAACACACCTGTAATAGTATGGCTACCGGCTACCACTCTAAATATTTAGAGATGGT
IleValProLeuAlaLysValAspCysThrAlaAsnThrAsnThrCysAsnLysTyrGlyValThrGlyTyrProThrLeuLysIlePheArgAspGly 109
397 GAAGAAGCGGTGCTTATGATGGCTAGGACTCCGATGGAATTTGTCAGCACTTGAAGAAACAGCAGGACAGCTTCAGTCTCTCAGGACTGAG
GluGluAlaGlyAlaTyrAspGlyProArgThrAlaAlaSerHisLeuLysLysGlnAlaGlyProAlaSerValProLeuArgThrGlu 142
496 GACGAATTAAGAAGTTCATTAGTGATAAAGATGCTCGGTGGTGGCTTTTTCAGGATTTATTCAGTGATGCCACTCCGAGTCTCTAAAGACGCC
AspGluPheLysLysPheIleSerAspLysAspAlaSerValValGlyPhePheArgAspLeuPheSerAspGlyHisSerGluPheLeuLysAlaAla 175
595 AGCAACTGGAGATAACTACGATTTGCCACACCAAGCTTGAATCTCTGGTGAAGGAGTACGATGATAATGGAAGGGGATTACTATATTCGTTCA
SerAsnLeuArgAspAsnTyrArgPheAlaHisThrAsnValGluSerLeuValLysGluTyrAspAspAsnGlyGluGlyIleThrIlePheArgPro 208
694 TTACATCTGGCTAAACAAATTTGAAGACAAATTTGGTGCATATACTGAAAGAAATGACCACTGCAATCAAGAAGTTTATTCAGGAAAGCATTTGGT
LeuHisLeuAlaLysPheGluAspLysIleValAlaTyrThrGlyLysMetThrSerGlyLysSerArgSerLeuPheArgLysAlaPheGly 241
793 CTCGTCTCATATGACAGAAATATAAAGATTTGATCAAGGCAAGGACTTACTACCGCTTACTATGTGGACTGGAATGAAAGAAATCTAAAGAT
LeuCysProHisMetThrGluAspAsnLysAspLeuIleGlnGlyLysAspLeuThrAlaTyrTyrAspValAspTyrGlyLysAsnThrLysGly 274
892 TCTAACTCTGGAGAAACAGGTCTATGATGGTGGCAAGACATCTTGTATGCTGGACACAACTCAACTTTGTCTAGTACGCTGAACCTTTAGC
SerAsnTyrTrpArgAsnArgValIleMetValAlaLysThrPheLeuAspAlaGlyHisLysLeuAsnPheAlaValAlaSerArgLysThrPheSer 307
991 CATGAATTGCTGACTTTGGCTTGAAGAAGCTACTGGAGAGCTTCTGTGGCTATCAGAACTGTAAGAGGAGAGGTTTGTCTAGCAGGAGGAG
HisGluLeuSerAspPheGlyLeuGluSerThrThrGlyGluIleProValValAlaIleArgThrAlaLysGlyGluLysPheValMetGlnGluGlu 340
1090 TTTCTCGGGATGGCAAGCTCTTGAGCGGTTCTGCAAGGAACTTTGTATGGCACTTGAAGAGATACCTGAAGTCTCAACTATCCCCGAGACCAAC
PheSerArgAspGlyLysAlaLeuGluArgPheLeuGlnGluTyrPheAspGlyAsnLeuLysArgTyrLeuLysSerGluProIleProGluThrAsn 373
1189 GAAGGACCTGTCAAGTGGTGGTAGCAGAGATTTTGTATGACATAGTGAATGCTGAAGACAAGGAGCTGCTGATCGAATTTATGCTCCTTGGTGGC
GluGlyProValLysValValValAlaGluSerPheAspAspIleValAsnAlaGluAspLysAspValIleLeuGluPheThrAlaProTrpCysGly 406
1288 CACTGTAAGAACCTGGAACCAAGTACAAAGAGCTGGGGGAGAACTCAGCAAGACCCAAATATTGTCTATGCAAGATGGATGGATCCACAGCTAT
HisCysLysAsnLeuGluProLysTyrLysGluLeuGlyGluLysLeuSerLysAspProAsnIleValIleAlaLysMetAspAlaThrAlaAsnAsp 439
1387 GTGCTCTTCCATATGAAGTCAAGGGTTTCTTACCTACTTCTCACCAGCCACAAGAAGCTAACTCAAGAAGTATGAAGTGGCGGTGAATTA
ValProSerProTyrGluValLysGlyPheProThrIleTyrPheSerProAlaAsnLysLysLeuThrProLysLysTyrGluGlyGlyArgGluLeu 472
1486 AATGATTTTATCAGCTATCTACACGAGAGCTACAAACCTCTATAATTCAAGAAGAAACCTTAAGAGAGAGAGGACACAGGAGGACTCTAA
AsnAspPheIleSerTyrLeuGlnArgGluAlaThrAsnProIleIleGlnGluGlyLysProLysLysLysLysAlaGlnGluAspLeu 504
1585 AGCAACAGCCAAATGACCACTTTATAAAGGACTCTTACACAGAGAAAGCAAAACCATCAGAGAGGAGAGATGGATATCTCTGAATCTGTATAA
1684 TTTTCTCTAACTGTTTCTAGCTGCAGTTTGTAAATACCAAGGACAGTTTATGTTTGTGGTTTGGGAGAAATATTTGTGTGGGAGAAATGTG
1783 GTAGGGGTGGGGGAAATGAGTGGGGGGTTATTTCTAATTTTTTTGTACATTGGACAGTGACATAAATGTGCCCTTTAAAAAATAAAAAA
1882 AAAAAAAAAAAAAAAAAAAAA 1902

```

PI-PLC I (24-133)	1	...	ASD	V	L	E	L	T	D	E	N	F	L	S	R	V	S	D	T	G	G	A	G	L	H	L	V	E	F	F	A	P	W	C	G	R	C	K	R	L	A	P	E	Y	E	50
PI-PLC I (374-484)		...	E	G	P	V	K	V	V	A	E	S	F	D	I	I	V	N	A	R	K	D	.	V	L	T	E	F	F	A	P	W	C	G	R	C	K	R	L	A	P	E	Y	E		
Spinach		K	A	S	E	A	V	K	E	V	D	V	N	S	G	W	K	E	F	V	L	Q	S	E	P	S	.	M	V	D	F	A	P	W	C	G	R	C	K	R	L	A	P	E	Y	
E. coli		...	M	S	D	K	I	I	H	L	T	D	S	F	D	T	D	V	L	K	A	D	G	A	.	I	L	V	D	F	A	P	W	C	G	R	C	K	R	L	A	P	E	Y		
G. naphthidii		...	A	T	V	K	V	D	N	S	N	F	Q	S	D	.	V	L	Q	.	S	S	E	P	.	V	V	D	F	A	P	W	C	G	R	C	K	R	L	A	P	E	Y			

PI-PLC I (24-133)	51	A	A	T	R	L	.	K	G	I	V	F	L	A	K	V	D	C	F	A	N	T	N	C	N	R	Y	G	V	T	G	F	T	I	F	R	G	E	E	A	G	A	.
PI-PLC I (374-484)		E	L	G	E	K	L	S	K	D	E	N	V	F	A	K	R	D	A	E	A	N	D	V	P	S	P	V	L	G	F	T	I	F	R	G	E	E	A	G	A	.	
Spinach		E	L	A	K	E	.	S	G	K	I	V	F	A	K	R	D	A	E	A	N	D	V	P	S	P	V	L	G	F	T	I	F	R	G	E	E	A	G	A	.		
E. coli		E	L	A	D	E	.	S	G	K	I	V	F	A	K	R	D	A	E	A	N	D	V	P	S	P	V	L	G	F	T	I	F	R	G	E	E	A	G	A	.		
G. naphthidii		E	L	A	T	E	M	.	A	G	Q	K	F	A	K	V	N	D	E	N	P	E	L	A	A	Q	F	G	F	R	S	I	P	T	L	M	F	R	G	E	L		

PI-PLC I (24-133)	101	V	D	G	P	R	T	A	D	G	I	V	S	H	L	K	K	Q	A	G	.
PI-PLC I (374-484)		V	E	G	G	K	E	L	N	D	F	I	S	V	L	R	E	A	.	.	
Spinach		G	D	V	S	K	Y	G	L	.	.	.	.	.	.	.	.	.	.		
E. coli		.	G	A	L	S	K	G	Q	L	.	K	E	F	L	D	A	N	L		
G. naphthidii		G	A	A	P	K	S	R	L	A	D	W	K	A	S	A	.	.	.		

Fig. 3 Alignment of rat PI-PLC I repeat with thioredoxins. The optimal alignment of PI-PLC repeat (24-133 and 374-484) with thioredoxins of type m spinach chloroplast thioredoxin²³, *E. coli* thioredoxin²⁴, and *Corynebacterina nephridii* thioredoxin²⁵ are shown. The amino acids comprising the conserved 14-membered disulphide ring are indicated by asterisks.

site. Overall, PI-PLC I repeat 24-133 exhibits 55.6% similarity with *Escherichia coli* thioredoxin, whereas PI-PLC I repeat 374-484 exhibits 51.5% similarity, an observation which could have important implications for either the regulation or catalytic activity of PI-PLC I.

E. coli thioredoxin has been reported to contain a labile phosphothiol linkage on Cys 32 and Cys 35, indicating that thioredoxin may function in phosphotransferase reactions^{18,19}. The high degree of similarity between the thioredoxins and the two repeats within PI-PLC I suggests that the conserved Cys-Gly-His-Cys domains within the repeats may play an important role in the phosphodiesterase hydrolysis of phosphoinositides, perhaps by forming a labile inositolide phosphothiol intermediate. Alternatively, interactions between PI-PLC I and regulatory proteins (for example GTP-binding proteins) may involve thiol-dependent redox reactions.

The tissue distribution of PI-PLC I messenger RNA was determined by Northern blot analysis of total RNA from a variety of rat tissues. PI-PLC I cDNA probe recognized a 2.0-kb mRNA in all the tissues sampled. The mRNA was most abundant in intestine and stomach and barely detectable in epidermis and skeletal muscle (Fig. 4a). Additional weakly hybridizing mRNA species were identified in several tissues, including adrenal, intestine, kidney and lung. These may represent mRNAs encoding related proteins (for example, other PI-PLC isoenzymes) or may be processing intermediates of the PI-PLC I transcript. To examine the frequency of PI-PLC I sequences within the genome, PLC-79 cDNA was used to probe Southern blots of

rat genomic DNA and washed to high stringency. Multiple restriction fragments totalling ~20 kb were identified for each enzymatic digest. These data suggest that PI-PLC I is encoded by a large gene or that it is a member of a conserved family of genes encoding related proteins.

Recently, Stahl *et al.* reported the DNA sequence of a 148K PI-PLC (PLC-148) from bovine brain¹¹ (previously termed PLC-II by Ryu *et al.*⁶). The presence of sequences similar to the non-catalytic domain of non-receptor tyrosine kinases and a newly described transforming gene *crk*²⁰ in PLC-148, led the authors to propose that PLC-148 may share a common regulatory mechanism with these enzymes¹¹. In comparing the sequences of PI-PLC I and PLC-148, we find that the two enzymes are quite dissimilar. Although PI-PLC I contains several short regions (10-25 amino acids) which share homology with PLC-148, including one which is similar to the 'A' domain found in tyrosine kinase (469-482), it lacks the 'B' and 'C' domains which are present in PLC-148 (ref. 11). Conversely, PLC-148 does not contain the conserved thioredoxin sequence found in PI-PLC I. This lack of sequence similarity is consistent with the absence of immunological cross-reactivity between the PI-PLC isoenzymes. Moreover, it suggests that the two enzymes do not share common regulatory sequences and may lack a common phosphodiesterase catalytic site.

The divergent sequence of the two PI-PLC isoenzymes is intriguing given our poor understanding of the role that the different PI-PLC isoenzymes play in agonist-induced phosphoinositide hydrolysis. An attractive hypothesis is that the

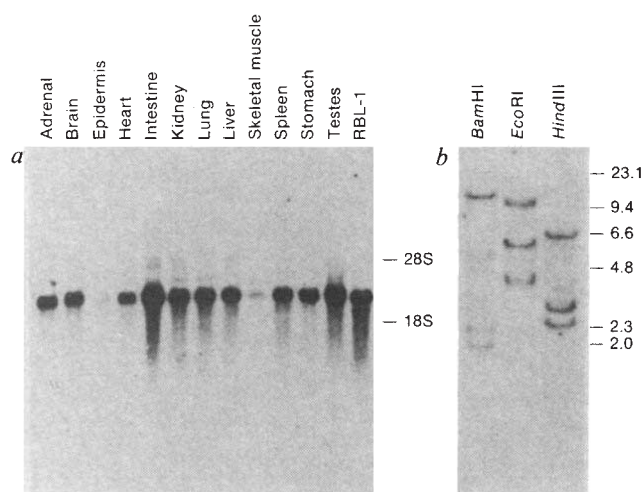


Fig. 4 Northern and Southern blot hybridization of PI-PLC I cDNA. *a*, Rat tissues were homogenized in 4-M guanidinium thiocyanate, 14% β -mercaptoethanol 10-mM Tris $\cdot$ HCl, pH = 7.5, and 5-mM EDTA²⁶. RNA was isolated by centrifugation through a bed of 5.7-M CsCl (100,000g, 18 h, 22 °C). The pellet was dissolved in 0.3% SDS (sodium dodecyl sulphate) and ethanol precipitated. Total RNA (15 μ g) from each tissue was electrophoresed, blotted and hybridized to ³²P-labelled PI-PLC I cDNA insert in 50% formamide, 5 $\times$ SSC, 1 $\times$ Denhardt's solution, 0.1% SDS and 100 μ g ml⁻¹ salmon sperm DNA for 20 h at 42 °C. The filters were washed in 1 $\times$ SSC, 0.2% SDS at 55 °C and autoradiographed. *b*, Southern blot hybridization of PI-PLC I cDNA to rat genomic DNA. Rat liver DNA (15 μ g) was digested with *Bam*HI, *Eco*RI and *Hind*III, electrophoresed in agarose, blotted to nitrocellulose filters and hybridized to ³²P-labelled PI-PLC I cDNA as described above. Filters were washed with 0.2 $\times$ SSPE (30 mM NaCl, 2 mM NaH₂PO₄, 0.2 mM EDTA; pH = 7.4), 0.2% SDS 68 °C and autoradiographed.

different isoenzymes couple to distinct classes of cell surface receptors. The tissue-specific distribution of PI-PLC isoenzymes would support such a model⁴⁻⁸. We also have evidence that PI-PLC I is solubilized as a complex containing PI-PLC I, a GTP-binding protein and vasopressin (V₁) receptors from rat liver membranes (N. Aiyar *et al.*, manuscript submitted), indicating that PI-PLC I is coupled to V₁ receptors in rat liver membranes. Expression of PI-PLC I in heterologous host systems should allow us to determine the functional specificity of PI-PLC I in the signal transduction process. Sequence analysis of other PI-PLC isoenzymes combined with site directed mutagenesis studies may aid in the identification of regulatory and catalytic sites for this important class of enzymes.

We thank Michael P. Angioli and Mark Strohsacker for technical assistance, Jon Marsh for oligonucleotide synthesis, and Dr Jeffrey Stadel for critical discussions. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries.

Received 1 April; accepted 10 May 1988.

1. Abdel-Latif, A. A. *Pharmac. Rev.* **38**, 227-272 (1986).
2. Berridge, M. J. A. *Rev. Biochem.* **56**, 159-193 (1987).
3. Takenawa, T. & Nagai, Y. *J. biol. Chem.* **256**, 6769-6775 (1981).
4. Hofmann, S. L. & Majerus, P. W. *J. biol. Chem.* **257**, 6461-6469 (1982).
5. Hakata, H., Kambayashi, J. & Kosaki, G. *J. Biochem.* **92**, 929-935 (1982).
6. Ryu, S. H., Cho, K. S., Lee, K.-Y., Suh, P.-G. & Rhee, S. G. *J. biol. Chem.* **262**, 12511-12518 (1987).
7. Ryu, S. H., Suh, P.-G., Cho, K. S., Lee, K.-Y. & Rhee, S. G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6649-6653 (1987).
8. Bennett, C. F. & Crooke, S. T. *J. biol. Chem.* **262**, 13789-13797 (1987).
9. Katan, M. & Parker, P. J. *Eur. J. Biochem.* **168**, 413-418 (1987).
10. Fox, J. A., Soliz, N. W. & Saltiel, A. R. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2663-2667 (1987).
11. Stahl, M. L., Ferez, C. R., Kelleher, K. L., Kriz, R. W. & Knopf, J. L. *Nature* **332**, 269-272 (1988).
12. Holmgren, A. A. *Rev. Biochem.* **54**, 237-271 (1985).
13. Synder, M., Sweetser, D., Young, R. A. & Davis, R. W. *Meth. Enzym.* **154** (in the press).
14. von Heijne, G. *J. molec. Biol.* **184**, 99-105 (1985).
15. von Heijne, G. *J. molec. Biol.* **189**, 239-242 (1986).

16. Kretsinger, R. H. *CRC Crit. Rev. Biochem.* **8**, 119-174 (1980).
17. Geisow, M. J., Fritsche, U., Hexham, J. M., Dash, B. & Johnson, T. *Nature* **320**, 636-638 (1986).
18. Pigiet, V. & Conley, R. R. *J. biol. Chem.* **253**, 1910-1920 (1978).
19. Conley, R. R. & Pigiet, V. *J. biol. Chem.* **253**, 5568-5572 (1978).
20. Mayer, B. J., Hamaguchi, M. & Hanafusa, H. *Nature* **332**, 272-275 (1988).
21. Sanger, F., Milken, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
22. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767-4771 (1987).
23. Maeda, K., Tsugita, A., Dalzoppo, D., Vilbois, F. & Schurmann, P. *Eur. J. Biochem.* **154**, 197-203 (1986).
24. Lim, C. J., Geraghty, D. & Fuchs, J. A. *J. Bacteriol.* **163**, 311-316 (1985).
25. Meng, M. & Hogenkamp, H. P. C. *J. biol. Chem.* **256**, 9174-9182 (1981).
26. Chirgwin, J., Przbyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).

Why ion pair reversal by protein engineering is unlikely to succeed

Jenn-Kang Hwang & Arieh Warshel

Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, USA

Genetic engineering is a powerful tool for exploring correlations between structure and function in proteins, but as yet we are unable to use it for effective protein design. One of the most interesting examples, which would seem to be obvious, is reversing the polarity of an ion pair. Changing a positively charged protein group, that provides a strong binding for negative substrates, to a negative group is expected to provide an effective binding site for a positively charged substrate. But several recent experiments on aspartate aminotransferase^{1,2}, trypsin³ and aspartate transcarbamoylase (Schachman, H. K. personal communication) have indicated that polarity reversal is not so successful. Here we argue that the same factors that make the enzyme an effective system for the (-+) pair will make it a much less effective system for the (+-) pair. We also point out that the unusually low effective dielectric constant ($\epsilon \approx 13$) for the (-+) interaction is due to its microenvironment and this will destabilize a (+-) arrangement having an entirely different dielectric constant ($\epsilon \approx 80$). The calculations presented here evaluate the energetics of ion pairs in protein active sites on a semiquantitative level. This is particularly important when dealing with strong, functionally important interactions that are difficult to evaluate with macroscopic models.

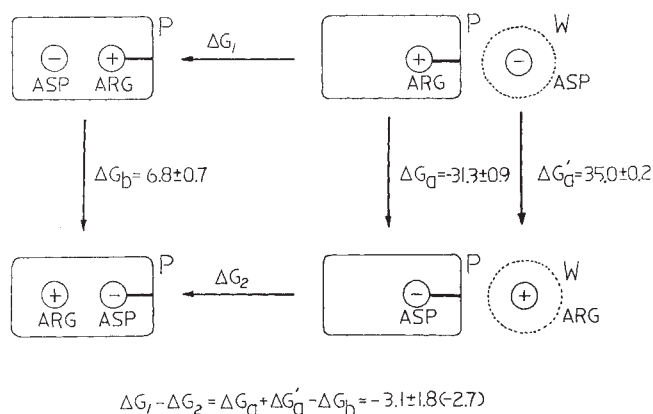
Perhaps the most striking example of the effect of polarity reversal is the Arg-292 $\rightarrow$ Asp mutation in aspartate aminotransferase (AAT)^{1,2} which results in an increase of ~ 7 kcal mol⁻¹ in the apparent activation energy, $\Delta G_{cat} - \Delta G_{bind}$, for its negatively charged substrate, L-Aspartate. Attempts to bind positively charged substrates to the mutant enzyme, and so to create a reversed ion pair, only reduce the apparent activation energy by $\sim 1-2$ kcal mol⁻¹ (as judged by comparing the binding of substrates with negative and positive side chains to the mutant enzyme). The change in binding energy upon polarity reversal is also very significant (more than 2.7 kcal mol⁻¹).

To examine polarity reversal we performed free-energy perturbation calculations on the AAT example. The calculations (summarized in Fig. 1) 'mutated' the AAT system, gradually changing Arg 292 to Asp and at the same time changing the bound substrate from Asp to Arg (the corresponding free energy is denoted by ΔG_b). In addition we completed the relevant thermodynamic cycle (see Fig. 1) by changing Arg 292 to Asp in the free enzyme (no substrate (ΔG_a) and in water $\Delta G'_a$). This cycle gives the difference between the binding free energy of an Asp substrate to the native enzyme and an Arg substrate to the mutant enzyme by:

$$\Delta G_1 - \Delta G_2 = \Delta G_a + \Delta G'_a - \Delta G_b \quad (1)$$

Related cycles have been used to compare the free energies of different charge configurations in water and proteins⁴⁻⁸. Other

Fig. 1 The free-energy cycle for polarity reversal in AAT. The figure analyses the binding of aspartate to the native enzyme and the binding of arginine to the Arg-292 → Asp mutant. The energies are given in kcal mol⁻¹ and the experimental value of ($\Delta G_1 - \Delta G_2$) is given in brackets. The notations P and W designate protein and water, respectively. The van der Waals interactions used for the calculations were described by $V_{ij} = A_i A_j r_{ij}^{-12} - B_i B_j r_{ij}^{-6}$, where V_{ij} (in kcal mol⁻¹) is the interaction between the *i* and *j* atom pair at a distance *r* (in Å), *A* and *B* are parameters given in kcal^{1/2} mol^{-1/2} Å⁶ and kcal^{1/2} mol^{-1/2} Å³, respectively. The *A* values are 1,313, 630 and 770 for the aspartate oxygens, the aspartate carbon and the arginine nitrogens respectively. These parameters reproduced the observed solvation free energy of ionized aspartic acid in solution and a solvation free energy of -45 kcal mol⁻¹ for the ionized arginine molecule in solution. Other parameters are identical to those used in refs 5-7.



recent works (such as refs 9 and 10) have also found free-energy calculation methods to be quite effective in studying related problems. The parameters used for the Asp and Arg residues are summarized in the caption of Fig. 1. The other parameters for the water molecules and the protein atoms, as well as the methods of calculations, are identical to those used in our earlier studies⁵⁻⁷. The only difference is associated with the treatment of long-range force. In the present work, which examines large changes in electrostatic forces, it was essential to extend the explicit solvent model around the active site (which was represented by a surface-constrained all-atom solvent (SCAAS) model^{5,8} of 11 Å radius) to a larger radius. This was done by surrounding the SCAAS region with a large grid (20 Å radius) of Langevin dipoles⁸. The region beyond this grid was represented by a continuum. Using a larger grid did not change our results significantly. Using only an 11 Å SCAAS solvent shell, surrounded by a continuum, gave errors of ~10 kcal mol⁻¹. Extending the SCAAS region to 20 Å, without using the Langevin grid, would have required the expensive use of thousands of explicit water molecules (rather than representing them by dipoles). The calculations were performed by our molecular modelling program MOLARIS. The initial coordinates of the native protein were taken from the X-ray structure of the cytosolic AAT protein¹¹ and those of the mutant were constructed from the native protein by a model-building program. Calculations with the coordinates of the mitochondrial AAT protein¹² were done on a more preliminary level.

The calculated results presented in Fig. 1 estimated a change in binding energy, $\Delta G_1 - \Delta G_2$, of $\sim -3.1 \pm 1.8$ kcal mol⁻¹ which compared well with the experimental value of -2.7 kcal mol⁻¹. The error range was estimated by repeating the calculations with different initial conditions (see ref. 5 for discussion of the problems associated with establishing a reliable error range). The agreement between the calculated and observed free-energy differences lends credibility to values of electrostatic free energies which were calculated rather than measured directly. Here we define the free energy of polarity reversal, ΔG_b , as the free energy associated with changing the (-+) to the (+-) configuration. The value of this free energy change (6.8 kcal mol⁻¹) is large (for example, in water the energy difference between the (+-) and (-+) ion pairs is zero) indicating that the native enzyme produces a very strong electrostatic complement to the polarity of the (-+) ion pair. For example, if the dielectric constant, ϵ , for the (+-) ion pair in the mutant enzyme is as large as 80, this will give ~ -1.4 kcal mol⁻¹ electrostatic free energy at 3 Å charge separation. In this case the (-+) ion pair, which is more stable than the (+-) ion pair by ~ 6.8 kcal mol⁻¹, will have an electrostatic free energy of $-1.4 - 6.8 = -8.2$ kcal mol⁻¹, which corresponds to an effective dielectric constant (see ref. 13) of ~ 13 ($\Delta G \approx -8.2 = -332/(\epsilon R) = -332/(13)$). Note also that the effective dielectric constant for a 3 Å separated ion pair in water is ~ 40 (ref. 13).

The concept of an effective dielectric constant can be quite misleading when applied to microscopic systems⁸, and different constants may be obtained for different properties¹³ and for different ion pairs (as is the case in this work). Thus it is more instructive to adopt a molecular point of view and to examine how a protein can stabilize ion pairs in a different way than water does. An extensive discussion of this issue has been given elsewhere^{8,13,14}, and the main conclusion is summarized in Fig. 2. As shown in Fig. 2, the protein can stabilize ion pairs by providing a cryptate-like environment of permanent dipoles (such as hydrogen bonds, carbonyls and fixed water molecules) which are polarized to stabilize the ion pair (see refs 14 and 15). A similar effect can be obtained by surrounding the ion pair with charges which are, in turn, stabilized by fixed dipoles (such as negatively charged groups near the positive side of an ion pair, as is the case with Asp-102 in trypsin⁸). Having a relatively fixed complementary electrostatic environment is one of the main tricks used by enzymes to stabilize ion pairs better than water can. Note in this respect that the water system has to 'pay' with water-water interaction for polarizing its dipoles towards an ion pair⁴.

With the concept of a prepolarized active site it appears (see Fig. 2) that the same environment which is folded to stabilize an ion pair with a given polarity, will 'repel' an ion pair of the opposite polarity. To demonstrate the equivalence between the situation drawn in Fig. 2 and the actual situation in AAT, we performed the calculations summarized in Fig. 3. In this calculation we evaluated the average electrostatic potential from the protein at the sites of the positive and negative charge centres. The active site potential (calculated in each case for the relaxed

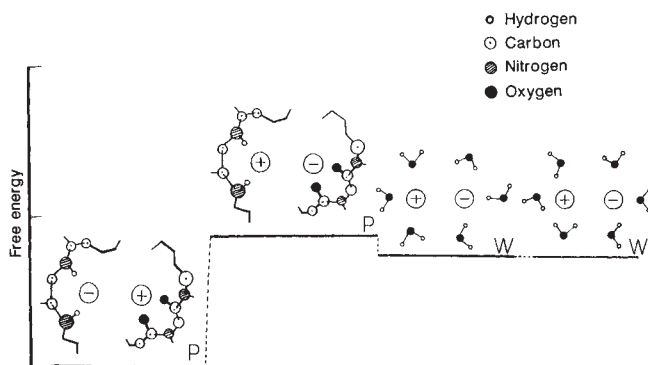


Fig. 2 Schematic description of the energetics of ion pair reversal in a hypothetical protein active site, and in water. The figure illustrates how the same surrounding protein environment that stabilizes a (-+) arrangement will destabilize a (+-) arrangement, while in water the energies of both configurations are identical.

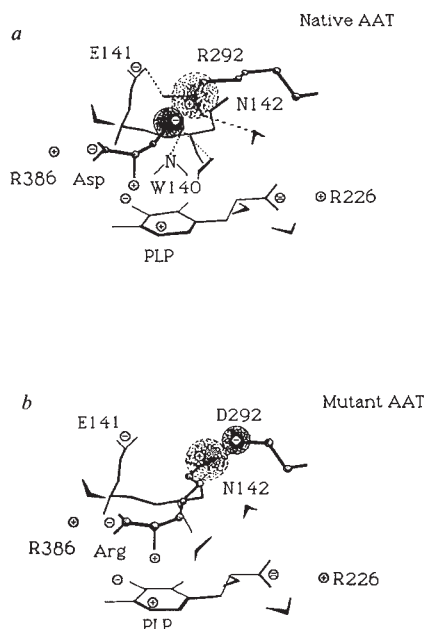


Fig. 3 The electrostatic microenvironment for the polarity reversal in AAT, and the average electrostatic potential from the surrounding environment on the relevant charges. The figure describes only some groups that give the large electrostatic contributions, but the actual calculations involve all the protein groups. In fact, what counts is the total potential and not the individual contributions which are strongly coupled (increasing one reduces the other). The potential is obtained after relaxing the environment at the presence of the given ion pair. Electric potentials are represented by spheres whose radii are proportional to the values of the potentials at the given sites. The calculated potential difference of the Asp-Arg ion pair in the native system is $184 \text{ kcal mol}^{-1}$ per electron charge and that of the corresponding ion pair in the mutant system is $165 \text{ kcal mol}^{-1}$ per electron charge. PLP designates the pyridoxal-5'-phosphate group.

configurations generated with the indicated charge distribution) thus favours the $(-+)$ over the $(+-)$ arrangement. The actual arrangement in AAT is quite different from the simple case of Fig. 2, as the protein uses charges as well as dipoles. But the effect is the same; the environment designed by nature to stabilize a given ion pair is not usually optimized for the inverted pair.

Recent papers^{2,3} have rationalized the problems associated with an effective ion pair reversal by arguing that one of the ions in the mutated pair is stabilized by some group of the protein (such as the H-bond between Asn-142 and Asp-292 in the mutated AAT (ref. 2), or the interaction between a main chain carbonyl and Lys-189 in the mutated trypsin³) and that this makes the ion inaccessible to its counter ion. Such an argument overlooks the fact that the same element that stabilizes an isolated ion would stabilize, rather than destabilize, the pair formed between this ion and counter ion. Thus, the N-H group in Fig. 2 stabilizes the $(-+)$ ion pair although it pulls the $(-)$ away from its $(+)$ counterpart. Similarly the N-H groups destabilizes the $(+-)$ pair although it pushes $(+)$ toward the $(-)$. It is important to realize that stable ion pairs do not involve strong attractive forces. For example, oppositely charged ions attract each other in vacuum more than in water, but their energy is much higher in vacuum than in water (see Fig. 6 of ref. 4).

In view of the arguments presented above we feel that polarity reversal would be an inherently ineffective genetic manipulation,

as long as the ion pair is more stable in its given environment than in water. Ion pairs which are not particularly stable lie in a local environment which is not prepolarized in any effective way. Reversing polarity in such cases might provide a somewhat more stable ion pair, but would still not create a very stable one. In this respect it is interesting to consider a recent work¹⁶ that describes reversals of ion pairs produced by genetic engineering. These studies involve ion pairs whose environments were not optimized by evolution and their successful reversal does not contradict our prediction. It would be interesting to attempt to design a properly polarized environment around a genetically engineered ion pair and then to examine the effect of polarity reversal.

An interesting exception to our prediction about polarity reversal may arise if an ionizable group can be kept inside a hydrophobic site. Designing such a site is not simple as ionized groups are unstable in hydrophobic environments. But if a covalent bond to the protein can keep such a group in its site despite its inherent instability (and if this group can be kept in its charged rather than neutral form), then it could provide an effective binding site for an oppositely charged group. Such a case, which is not common in proteins, may provide a system for an effective polarity reversal.

Finally we would like to point out that the AAT system can serve as a powerful test of various approaches for the calculation of electrostatic energies in proteins. That is, cases of weak interactions between surface groups in proteins can be modelled by representing the solvent with macroscopic models with large dielectric constants¹⁷⁻¹⁹ and even by Coulomb's law type models^{8,13}. The evaluation of strong short-range interactions might require more microscopic models^{4-8,20}. The AAT case is an excellent example, as it involves a very different effective dielectric constant for the $(+-)$ and $(-+)$ cases and as such cannot be accurately treated by a Coulomb's law type model. As shown in the present work, it is indeed possible to handle this system using free energy perturbation approaches. Such approaches, however, are very time consuming and make it difficult to extract a simple physical insight. Thus it would be useful if one could reproduce the energetics of the AAT case by semimicroscopic⁸, or macroscopic¹⁷⁻¹⁹ grid models.

This work was supported by the NSF and the Office of Naval Research. The authors are grateful to Professor Arthur Arnone and Professor Johan N. Jansonius for the coordinates of the proteins and to Professor Jack F. Kirsch for making his work available before publication. We are grateful to Alliant Inc. for a donation of an award that will be given to the first person who succeeds in an effective polarity reversal (provided the native ion pair is significantly stabilized by the protein environment).

Received 8 March; accepted 7 June 1988.

1. Cronin, C. N., Malcolm, B. A. & Kirsch, J. F. *J. Am. chem. Soc.* **109**, 2222-2223 (1987).
2. Cronin, C. N. & Kirsch, J. F. *Biochemistry* (manuscript submitted).
3. Graf, L. *et al. Biochemistry* **26**, 2616-2638 (1987).
4. Warshel, A. *Biochemistry* **20**, 3167-3177 (1981).
5. Warshel, A., Sussman, F. & King, G. *Biochemistry* **25**, 8368-8372 (1986).
6. Warshel, A. & Sussman, F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3806-3810 (1986).
7. Hwang, J.-K. & Warshel, A. *Biochemistry* **26**, 2669-2673 (1987).
8. Warshel, A. & Russell, S. T. *Q. Rev. Biophys.* **17**, 283-422 (1984).
9. Rao, S. N., Singh, U. C., Bash, P. A. & Kollman, P. A. *Nature* **328**, 551-554 (1987).
10. Mezei, M., Mehrotra, P. K. & Beveridge, D. L. *J. Am. chem. Soc.* **107**, 2239-2245 (1985).
11. Arnone, A. *et al. in Molecular Structure and Biological Activity* (eds Griffen, J. F. & Duax, W. L.) 57-74 (Elsevier, New York, 1982).
12. Kirsch, J. F. *et al. J. mol. Biol.* **174**, 497-525 (1984).
13. Warshel, A., Russell, S. T. & Churg, A. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4785-4789 (1984).
14. Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5250-5254 (1978).
15. Quiocho, F. A., Sack, J. S. & Vyas, N. K. *Nature* **329**, 561-564 (1987).
16. Wells, J. A., Powers, D. B., Bott, R. R., Craycar, T. P. & Estell, D. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1219-1223 (1987).
17. Sternberg, M. J. E., Hayes, F. R. F., Russell, A. J., Thomas, P. G. & Fersht, A. R. *Nature* **305**, 86-88 (1987).
18. Russell, A. J. & Fersht, A. R. *Nature* **328**, 496-500 (1987).
19. Gilson, M. K. & Honig, B. H. *Nature* **330**, 84-86 (1987).
20. Warshel, A. *Nature* **330**, 15-16 (1987).

Solid phase extraction fundamentals

B. Tippins

Sample preparation columns based on the principles of solid phase extraction are fast becoming indispensable tools in many types of research.

SOLID phase extraction is a sample preparation technique that utilizes disposable extraction columns which contain small amounts of bonded silica materials to extract rapidly and quantitatively specific compounds of interest from complex samples. Solid phase extraction is used to prepare samples for analysis by techniques such as high performance liquid chromatography, gas chromatography, nuclear magnetic resonance and radioimmunoassay because it provides a clean, concentrated extract which simplifies subsequent analysis by removing interfering or cross-reacting components. It is used in a wide variety of applications, including the extraction of pollutants from soil and water¹⁻³, drugs of abuse from body fluids^{4,6}, and peptides from cell culture or serum⁷⁻⁹. Extractions may take as little as five minutes from start to finish, and with proper development may lead to 90 per cent recovery of the compound of interest from the sample matrix.

Many types of bonded silica materials are available, and each one has a slightly different selectivity. Bonded silica columns can be used with any solvent, and the short extraction time required allows them to be used with either strongly acidic or basic solutions. This combination of advantages makes solid phase extraction a powerful and versatile sample preparation technique.

The method development process identifies which combination of bonded silica phase and solvent extraction conditions will yield the cleanest extract with maximal recovery. Bonded silica materials typically fall into three major categories based on the type of chemical interaction between the compound of interest and the bonded silica material: non-polar, polar or ion exchange. The first step in method development is to select one of these extraction techniques based on the chemical characteristics of the compound of interest and the sample matrix. Each type of extraction is performed in a different way, with its own set of rules which must be followed to obtain the best results.

Non-polar extraction

Non-polar extraction columns contain silica particles bonded to octadecyl, octyl, ethyl, phenyl or other hydrophobic functional groups. These columns are used to extract hydrophobic compounds from aqueous solutions such as body fluids, wastewater samples or wine samples.

Non-polar extraction columns are typically prepared for an extraction process by rinsing with methanol followed by water. The hydrophobic compound of interest is retained by a hydrophobic interaction with the solid phase, and an aqueous rinse solution is passed through the column to remove the undesired polar sample components. The compound of interest is then eluted from the column with a small volume of an organic solvent such as methanol or acetonitrile.

If there are ionic groups present on the compound of interest, retention may

groups such as diol, cyanopropyl or aminopropyl. They are used to extract polar or hydrogen-bonding compounds from non-aqueous, non-polar samples such as vegetable oil or hexane.

In a polar extraction, the column is prepared by rinsing it with a non-polar solvent such as hexane or chloroform before applying the sample. The compound of interest is retained by hydrogen bonding with the solid phase. A non-polar solvent is used to rinse the undesired non-polar matrix components from the column. The compound of interest then is

Table 1 Extraction column selection guide

	Extraction Column	Analyte Functional Groups	Matrix	Typical Elution Solvents	Typical Applications
Non-polar Extraction	C18 – Octadecyl C8 – Octyl C2 – Ethyl CH – Cyclohexyl PH – Phenyl CN – Cyanopropyl CN – Cyanopropyl (endcapped)	<i>Hydrophobic Groups:</i> Aromatic rings Alkyl chains	<i>Aqueous:</i> Water Buffers Biological Fluids	Methanol Acetonitrile Ethyl Acetate Chloroform Acidic Methanol Hexane	Drugs of Abuse Peptides Pesticides Therapeutic Drug Monitoring
Polar Extraction	CN – Cyanopropyl 2OH – Diol SI – Silica NH ₂ – Aminopropyl PSA – Primary/-Secondary Amine	<i>Hydrophilic Groups:</i> Hydroxyls Amines Heteroatoms (S, O, N)	<i>Non-Polar:</i> Hexane Oils Chloroform Lipids	Methanol Isopropanol Acetone	Vitamin D Metabolites Lipid Separations Oil Additives Carbohydrates Phenols
Cation Exchange Extraction	SCX – Benzenesulfonic Acid (Strong) PRS – Propylsulfonic Acid (Strong) CBA – Carboxylic Acid (Weak)	<i>Cations:</i> Amines Pyrimidines	<i>Aqueous:</i> Water Acidic Buffers Biological Fluids	Alkaline Buffer High Ionic Strength Buffer	Catecholamines Herbicides Pharmaceuticals
Anion Exchange Extraction	SAX – Quaternary Amine (Strong) PSA – Primary/-Secondary Amine NH ₂ – Aminopropyl (Weak) DEA – Diethylaminopropyl (Weak)	<i>Anions:</i> Carboxylic Acids Sulfonic Acids Phosphates	<i>Aqueous:</i> Water Alkaline Buffers Biological Fluids	Acidic Buffer High Ionic Strength Buffer	Organic Acids Vitamins Fatty Acids Phosphates

be improved by adjusting the pH of the sample to neutralize the charge, making the compound as hydrophobic as possible. The rinse solution should be maintained at this adjusted pH to reduce the likelihood of losing the compound in the rinse step. It is not uncommon for compounds which contain amine groups to be strongly retained by both the primary hydrophobic interaction and a strong secondary ionic interaction with residual unbonded silanol groups on the silica particle. This problem is solved easily by using a solution composed of a 95:5 ratio by volume of methanol to hydrochloric acid as the elution solvent. This improves recovery by disrupting both interactions simultaneously.

Polar extraction

Polar extraction columns contain either unbonded silica, or silica bonded with polar or hydrogen-bonding functional

eluted with a more polar solvent such as methanol, isopropanol or a combination of non-polar and polar solvents such as hexane/isopropanol.

Water remaining in the sample will compete for the hydrogen bonding sites on the column, dramatically reducing the column's capacity to bind the compound of interest. For best results, residual water must be removed from the sample before performing a polar extraction.

The polar extraction method is selective enough to separate metabolites which differ by only one hydroxyl group. This is done by sequential elution with stepwise stronger elution solvents. For example, the more non-polar compound may be eluted with a solution containing a 95:5 ratio by volume of hexane to isopropanol, and the more polar compound may be eluted with a solution of the same chemicals in a ratio of 85:15.

Ion exchange extraction

Ion exchange columns can be either cationic or anionic. It is probably the most selective extraction mechanism, and yields the cleanest extracts. The pH of the sample and all solutions should be carefully controlled; a sample with high ionic strength (over 0.2 M) should be diluted with water to enhance retention. The flow rate during sample application and elution must be relatively low (about 1-2 ml min⁻¹) for the ion exchange interactions to occur.

Cation exchange columns contain silica bonded with sulfonic or carboxylic acid functional groups. These are used to extract amine-containing compounds from aqueous or non-aqueous solutions.

The columns are conditioned with methanol followed by a buffer. The pH of the conditioning buffer, the sample and the rinse solution should be neutral, because at this pH, the amines on the compound of interest carry a positive charge and the acidic functional groups on the silica carry a negative charge, allowing retention by a strong ionic interaction. A rinse buffer can be used to remove any non-polar compounds from the column. The compound of interest may then be eluted from the column using an acidic wash to neutralize the bonded silica, or by using a basic wash to neutralize the charge on the compound of interest. Use of a 50 per cent methanol solution in the elution wash enhances recovery by disrupting secondary hydrophobic interactions.

Anion exchange columns contain silica bonded with amine functional groups such as aminopropyl or quaternary amines, and are used to extract compounds which contain carboxylic acids or sulphonic acids. Again, the pH of the conditioning buffer, the sample and the column rinses should be neutral to ensure ionization of both the acidic compound of interest and the basic extraction column. Elution can be done by neutralizing the charge on either one.

Solid phase extraction provides a rapid and effective way to clean up and concentrate a variety of compounds for analysis. But developing a reliable extraction method depends upon understanding the properties of bonded silica materials. □

Barbara L. Tippins is at Analytichem International, 24201 Frampton Avenue, Harbor City, California 90710, USA. For more information, fill in reader service number 100.

- Ozretich, R.J. & Schroeder, W.P. *Analyt. Chem.* **58**, 2041-2048 (1986).
- Rostad, C.E., Pereira, W.E., Kutchell, S.M. *Analyt. Chem.* **58**, 2856-2860 (1984).
- Chladek, E. & Marano, R.S. *J. chromatogr. Sci.* **22** 313-320 (1984).
- Werner, M., Hertzman, M., Pauley, C.J. *Clin. Chem.* **32**, 1921-1924 (1986).
- Ford, B., Vine, J., Watson, T.R. *J. Analyt. Tox.* **7**, 116-118 (1983).
- Wilson, I.D., Ismail, I.M. *J. Pharm. biomed. Anal.* **4**, 663-665 (1986).
- Nussberger, J., Brunner, D.B., Waeber, B., Brunner, H.R. *Hypertension* **7**, supplement 1, 1-7 (1985).
- Burnett, J.D., Jr. *et al. Science* **231** 1145-1147 (1985).
- Ohara, J., Lahet, S., Inman, J. Paul, W.E. *J. Immun.* **135**, 2518-2523 (1985).

Analytical techniques

This week's review covers products for all types of analyses, from electrophoresis equipment to chromatography and spectroscopy instruments.

SPECTROVISION has just announced the availability of a new **scanning fluorescence spectrophotometer** system which it says was designed to yield excellent sensitivity and still be easy to use (*Reader Service No. 101*). The \$7,495 (US) system includes the SpectroVision model FD-300 spectrophotometer optical module and the FD-10 scan controller. It can be fitted with a one-position cell holder, a four-position thermostatted cell holder or an HPLC



Scanning fluorescence from SpectroVision.

flow cell. The spectrophotometer has a xenon flash lamp that SpectroVision says provides continuous excitation between 200 nm and 800 nm, plus a photomultiplier, and holographic excitation and emission gratings in its monochromators. The scan controller can be interfaced to a microcomputer for automated analyses, and the data output can be sent to a microcomputer for archival purposes and data manipulation.

A **laboratory reactor system** for chemical process development is just out from Guided Wave International AB (*Reader Service No. 102*). The Chemtech Reactor system was introduced at the Achema meeting in Frankfurt three months ago, and is intended for kinetic and mass transfer studies involving gases, liquids and solids in dynamic or stationary experiments. It is a modular and flexible reactor system with computerized control for the automation of every step, from dosing to cleaning. The application range of the Chemtech Reactor includes heterogeneous catalysis, cellulose processing and polymerization. The reactor has several connections in its lid and walls for use with a number of different sensors. The Guided Wave Analyzer, sold separately, can be used with the Chemtech Reactor for *in situ* sensing of chemical composition.

Ionic measurement

Magical is the name of the latest **image analysis system** from Joyce-Loebl Ltd, for the quantitative study of the spatial distri-

butions of Ca²⁺, Na⁺, pH and other ions within living cells, and their temporal changes in response to stimuli (*Reader Service No. 103*). With Magical, sequences of images are captured at normal television video rate, at wavelengths characteristic of the fluorescent dye being used. The fluorochromes Fura-2 and Indo-1 are used for calcium measurement, while DCH and BCECF show pH distributions and FCYR-2 shows sodium ions within cells. A filter changer alternates between wavelengths. The £30,000 (UK) system accommodates most inverted microscopes, and uses an intensifying camera which Joyce-Loebl says has a sensitivity of greater than 10⁻¹ lx. Magical captures images in user-selectable sizes from 64 × 64 to 512 × 512 pixels, and digitizes them into 64 or 256 grey levels. Up to 32 megabytes of image storage are provided within the image processing and analysis unit, which ratios corresponding pairs of images from alternate wavelengths. Magical's specialized software



Joyce-Loebl's Magical system.

provides a variety of functions, from profiling to graphing measurements of specified cell areas against time.

Russell pH Ltd has launched a kit for **measuring pH** in low ionic strength water (*Reader Service No. 104*). Russell pH says the £85 (UK) kit can be used for pH measurement of "pure water" samples — water in which the total amount of acid or

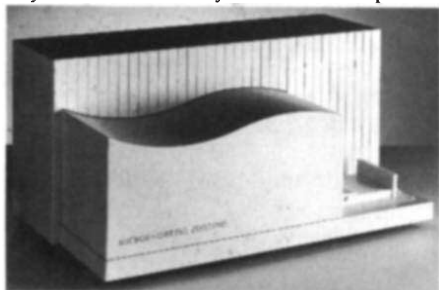


The kit for pure water studies.

base is very small and in which there is a low level of dissolved salts — including distilled waters, deionized waters, process waters, well and some surface waters, treated effluent waters and boiler feed waters. The kit includes electrode, buffers, inter-connecting cable and instructions for use. The kit's electrode is a combination pH electrode that has a low-resistance pH membrane. The annular design of the reference junction gives low junction resistance and eliminates problems associated with generated junction potentials, says the company.

Electrophoresis extras

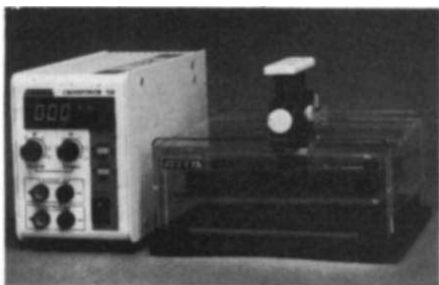
Microphoretic Systems has brought **automated capillary electrophoresis** to routine laboratory work with the Microphore 1000 instrument (*Reader Service No. 105*). The \$50,000 (US) Microphore 1000 has a patented multichannel UV/fluorescence detection system, and an automated sampling system based on standard 96-well microplates which can introduce samples by either controlled vacuum or electromigration techniques. Both the changing and replenishment of multiple electrolytes and the purging of the capillary between analyses are computer-



The new wave in capillary electrophoresis.

controlled. Microphoretic Systems says the instrument can be used for isoelectric focusing, isotachopheresis and micellar electrophoresis, and separations based on molecular size — such as restriction fragment analyses and SDS-PAGE — can be accomplished using gel-filled capillaries.

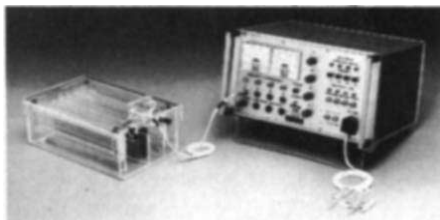
Genetic Research Instrumentation Ltd says its Horizblot **semi-dry electrophoresis blotter** can transfer DNA, RNA and protein samples from gels onto transfer membranes without denaturation from heat buildup (*Reader Service No. 106*). The £520 (UK) Horizblot can accept up to six gel or membrane matrixes simulta-



Transfer without denaturation with Horizblot.

neously. The company says the Horizblot's flat solid carbon electrodes ensure even pressure and contact over the whole of the gel surface, and prevent air bubble migration. The Horizblot is supplied with a full safety lid, adjustable electrodes and safety input leads. Genetic Research Instrumentation recommends the use of its Cross Power 150 electrophoresis power supply, available in 15 × 15 cm and 20 × 20 cm sizes, for use with the Horizblot.

A complete system for **qualitative and quantitative electrophoresis** — including power pack, tank with 120 × 220 mm



Accurate's equipment for electrophoresis.

working surface, electrodes and ancillary equipment — is now available from Accurate Chemical and Scientific Corporation (*Reader Service No. 107*). Accurate says the system is suitable for research and for routine immunology and biochemistry work, and is based on the apparatus developed by C.B. Laurell. It is intended for use in protein and immunoelectrophoresis gel techniques on glass, agar, paper or other media. The \$1,895 (US) system features solid-state electronics, and is offered with a complete line of accessories and spare parts.

Spectroscopy angles

Surface Analysis Technology plc now offers a service for **resonant ionization mass spectrometry**, a surface analysis technique which yields highly specific molecular information and low detection limits with a high degree of spatial resolution (*Reader Service No. 108*). The resonant ionization MS analysis begins with the laser ablation of a few millionths of a square millimetre of the sample surface, as in laser ionization mass analysis. But in resonant ionization, a second tunable pulsed dye laser is fired orthogonally into the plasma created by the first laser, resulting in greater ionization of the sample, and pushing the detection sensitivity of the technique down to parts per trillion. The ions are subsequently analysed in a time-of-flight mass spectrometer. Surface Analysis Technology says the system they use for the service provides a specificity of ionization that allows the investigation of large organic molecules, including plastics, polymers and biological molecules.

The newest **spectroscopy amplifier** from EG & G Instruments offers four automatic features that it says eliminate the

need for screwdriver adjustments: automatic pole-zero adjustment, automatic noise discriminators, automatic BLR rate and automatic compensation for reset recovery (*Reader Service No. 109*). The model 672 spectroscopy amplifier includes

ADVERTISEMENTS



Scientific Software Your search ends here...

- Statistics and forecasting
- Scientific data plotting
- Maths and operations research
- Technical word processing
- Data acquisition and analysis
- Control and simulation
- Equation solvers... and more

Science and Engineering
catalogue available

The Core Store Limited
59 Station Road, Northwich,
Cheshire CW9 5LT.
Tel: 0606 45420

Reader Service No.58

BIANORM EQUILIBRIUM DIALYSIS SYSTEM

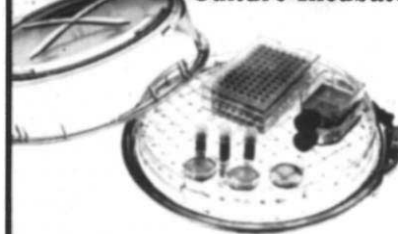


The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

BIANORM
Gerrit
P.O. Box 850126
D-4001 Munich 65,
FRG

Reader Service No.49

Economical Modular Tissue Culture Incubator



- Safe for AIDS research • Reliable, for In-Vitro fertilization
 - Versatile enough for thousands of lab uses
- Easy to use — Just flush with gas mixture, seal and place at approp. temp. Perform multiple experiments simultaneously w/out cross contamination. Thousands in use worldwide.

For information contact: Flow Laboratories, Ltd. or,
Billups-Rothenberg, Inc. • P.O. Box 977
Del Mar, CA 92014-0977 USA • (619) 755-3309

Reader Service No.25

a choice of triangular or Gaussian filters, for optimum resolution by effectively doubling the number of time constants available. It also features a differential input for the reduction of ground loop noise. EG & G says the amplifier can be used with germanium, silicon surface-barrier, silicon-lithium, and scintillation detectors, as well as with proportional counters.

Chromatography trends

Spectra-Physics has added the model SP8775 **liquid chromatography autosampler** to its range of modular liquid chromatography equipment (*Reader Service No. 110*). The company says the SP8775 was designed for reliability and simplicity of operation, for routine use. The autosampler connects to any pump or data system, and has an automatic flush function after each injection to prevent sample carryover. It has four removable trays which accommodate 80 samples plus a priority vial, and accept either crimp-top or screw-cap vials. The keyboard locks to protect against accidental interruption of a sampling run, and autosampler functions are indicated on the display panel during all stages of operation. The SP8775



The LC autosampler from Spectra-Physics.

can be configured for micro-liquid chromatography as well as large-volume injections, and an optional communications board allows it to be interfaced into laboratory networks. Spectra-Physics has given the \$6,500 (US) autosampler a five-year warranty.

The **MUST line of HPLC stream-switchers** from Spark Holland has been expanded. The MUST family now offers a wide choice of stream-switching functions, ranging from high-pressure applications such as heart-cut, backflush and boxcar applications, to low-pressure solvent selectors (*Reader Service No. 111*). The units may be remote-controlled, or be monitored by an integrated micro-processor timer. Dual- and single-valve units are available. Spark Holland says the MUST line is the perfect solution for on-

line sample clean-up and pre-concentration problems. The effectiveness of the stream-switchers is enhanced by disposable fluoropolymer pre-column cartridges, also available from Spark Holland.

Life Science Laboratories Ltd has recently introduced a new **gradient solvent**



Makes the gradient for SFC.

delivery system for supercritical fluid chromatography (SFC) (*Reader Service No. 112*). The solvent delivery system combines a 10,000 p.s.i. syringe pump with a computer-based SFC density gradient controller. Life Science Labs says the system is an affordable way to adapt almost any gas chromatograph for the SFC analysis of difficult samples not suitable for liquid chromatography or gas chromatography. The pump delivers flow rates of between 0.2 and 600 $\mu\text{l min}^{-1}$, and pump refill and repressurization is accomplished in minutes, says the company. Three gradient modes are possible: linear pressure, linear density-specified, and curved (asymptotic) density-specified. The software supplied with the system includes density-to-pressure conversion files for some common SFC solvents such as CO_2 , NH_3 and alkanes. The SFC pump controller interface may be installed in any IBM-PC or compatible.

Dionex is trying to make **routine chromatography** easier with the introduc-



HPLC customizing kits from Dionex.

tion of its **Dionex Analyzer Pacs (DAPs)** (*Reader Service No. 113*). Each DAP includes menu-driven software, instructions and components to customize any Dionex chromatograph and data station for a specific application. The program disk provided with each DAP package automatically programs the Dionex AutoIon 450 data station to set all instrument parameters and manage the collection, computation and printout of results. The DAPs also include a Dionex guard and analytical column designed for the application, and Dionex OnGuard sample pretreatment cartridges to extend the life of the analytical column. Dionex is developing DAPs for a host of applications, including packages for the analysis of ions in environmental samples.

Hybrid techniques

Hewlett-Packard has just introduced a **particle-beam liquid chromatography/mass spectrometer** that it says routinely generates library-matchable electron-impact spectra of nonvolatile and thermally labile compounds, to make LC/MS just as viable a technique as GC/MS (*Reader Service No. 114*). The complete \$135,000–200,000 (US) system consists of the particle-beam interface mounted on the HP5988A MS, a choice of either an integrated or a modular LC, and either a single-instrument or a multi-instrument data system. A gas chromatograph can be added to the configuration so the hybrid instrument can be switched from LC/MS to GC/MS capability. Options for the system include thermospray, fast-atom bombardment and a direct insertion probe.

Benchtop GC/MS is now available from VG Instruments with the introduction of the TRIO-1 (*Reader Service No. 115*). The £75,000 (UK) TRIO-1 features a differentially turbomolecular pumped, high-transmission, 12-mm in diameter by 200 mm in length quadrupole analyser with VG Instruments' contamination rejection system of pre- and post-filters. VG Instruments says the TRIO-1's vacuum housings are manufactured from extruded aluminium to maintain ion optical integrity. The ion source is constructed totally of metal, with no ceramic dowels or spacers, with their attendant fragility. The TRIO-1 takes advantage of the high speed of transputers for data processing and communication of data to the host computer. The LabBase software performs acquisition, mass chromatogram generation, automatic quantitation, spectral enhancement, library searching and plotting. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.